Numerical methods and models in market risk and financial valuations area

Jose Arturo Infante Acevedo

THÈSE

présentée pour l'obtention du titre de

Docteur de l'Université Paris-Est

Spécialité : Mathématiques Appliquées

par **José Arturo Infante Acevedo**

Ecole Doctorale : Mathématiques et Sciences et Technologies de l'Information et de la Communication

Méthodes et modèles numériques appliqués aux risques du marché et à l'évaluation financière

Soutenue le 09/12/2013
devant le jury composé de :

Rapporteurs :	Frédéric Abergel
	Yves Achdou
Président :	Bernard Lapeyre
Examinateur :	Mohamed Baccouche
Directeurs de thèse :	Aurélien Alfonsi
	Tony Lelièvre

Méthodes et modèles numériques appliqués aux risques du marché et à l'évaluation financière

Ce travail de thèse aborde deux sujets : (i) L'utilisation d'une nouvelle méthode numérique pour l'évaluation des options sur un panier d'actifs, (ii) Le risque de liquidité, la modélisation du carnet d'ordres et la microstructure de marché.

Premier thème : Un algorithme glouton et ses applications pour résoudre des équations aux dérivées partielles

Beaucoup de problèmes d'intérêt dans différents domaines (sciences des matériaux, finance, etc) font intervenir des équations aux dérivées partielles (EDP) en grande dimension. L'exemple typique en finance est l'évaluation d'une option sur un panier d'actifs, laquelle peut être obtenue en résolvant l'EDP de Black-Scholes ayant comme dimension le nombre d'actifs considérés. Nous proposons d'étudier un algorithme qui a été proposé et étudié récemment dans [ACKM06, BLM09] pour résoudre des problèmes en grande dimension et essayer de contourner la malédiction de la dimension. L'idée est de représenter la solution comme une somme de produits tensoriels et de calculer itérativement les termes de cette somme en utilisant un algorithme glouton. La résolution des EDP en grande dimension est fortement liée à la représentation des fonctions en grande dimension. Dans le Chapitre 1, nous décrivons différentes approches pour représenter des fonctions en grande dimension et nous introduisons les problèmes en grande dimension en finance qui sont traités dans ce travail de thèse.

La méthode sélectionnée dans ce manuscrit est une méthode d'approximation non-linéaire appelée Proper Generalized Decomposition (PGD). Le Chapitre 2 montre l'application de cette méthode pour l'approximation de la solution d'une EDP linéaire (le problème de Poisson) et pour l'approximation d'une fonction de carré intégrable par une somme des produits tensoriels. Un étude numérique de ce dernier problème est présenté dans le Chapitre 3. Le problème de Poisson et celui de l'approximation d'une fonction de carré intégrable serviront de base dans le Chapitre 4 pour résoudre l'équation de Black-Scholes en utilisant l'approche PGD. Dans des exemples numériques, nous avons obtenu des résultats jusqu'en dimension 10.

Outre l'approximation de la solution de l'équation de Black-Scholes, nous proposons une méthode de réduction de variance des méthodes Monte Carlo classiques pour évaluer des options financières.

Second thème : Risque de liquidité, modélisation du carnet d'ordres, microstructure de marché

Le risque de liquidité et la microstructure de marché sont devenus des sujets très importants dans les mathématiques financières. La dérégulation des marchés financiers et la compétition entre eux pour attirer plus d'investisseurs constituent une des raisons possibles. Les règles de cotation sont en train de changer et, en général, plus d'information est disponible. En particulier, il est possible de savoir à chaque instant le nombre d'ordres en attente pour certains actifs et d'avoir un historique de toutes les transactions passées. Dans ce travail, nous étudions comment utiliser cette information pour exécuter de facon optimale la vente ou l'achat des ordres. Ceci est lié au comportement des traders qui veulent minimiser leurs coûts de transaction.

La structure du carnet d'ordres (Limit Order Book) est très complexe. Les ordres peuvent seulement être placés dans une grille des prix. A chaque instant, le nombre d'ordres en attente d'achat (ou vente) pour chaque prix est enregistré. Pour un prix donné, quand deux ordres se correspondent, ils sont exécutés selon une règle First In First Out. Ainsi, à cause de cette complexité, un modèle exhaustif du carnet d'ordres peut ne pas nous amener à un modèle où, par exemple, il pourrait être difficile de tirer des conclusions sur la stratégie optimale du trader. Nous devons donc proposer des modèles qui puissent capturer les caractéristiques les plus importantes de la structure du carnet d'ordres tout en restant possible d'obtenir des résultats analytiques.

Dans [AFS10], Alfonsi, Fruth et Schied ont proposé un modèle simple du carnet d'ordres. Dans ce modèle, il est possible de trouver explicitement la stratégie optimale pour acheter (ou vendre) une quantité donnée d'actions avant une maturité. L'idée est de diviser l'ordre d'achat (ou de vente) dans d'autres ordres plus petits afin de trouver l'équilibre entre l'acquisition des nouveaux ordres et leur prix.

Ce travail de thèse se concentre sur une extension du modèle du carnet d'ordres introduit par Alfonsi, Fruth et Schied. Ici, l'originalité est de permettre à la profondeur du carnet d'ordres de dépendre du temps, ce qui représente une nouvelle caractéristique du carnet d'ordres qui a été illustré par [JJ88, GM92, HH95, KW96]. Dans ce cadre, nous résolvons le problème de l'exécution optimale pour des stratégies discrétes et continues. Ceci nous donne, en particulier, des conditions suffisantes pour exclure les manipulations des prix au sens de Huberman et Stanzl [HS04] ou de Transaction-Triggered Price Manipulation (voir Alfonsi, Schied et Slynko). Ces conditions nous donnent des intuitions qualitatives sur la manière dont les teneurs de marché (market makers) peuvent créer ou pas des manipulations des prix.

Numerical methods and models in market risk and financial valuations area

This work is organized in two themes : (i) A novel numerical method to price options on many assets, (ii) The liquidity risk, the limit order book modeling and the market microstructure.

First theme : Greedy algorithms and applications for solving partial differential equations in high dimension

Many problems of interest for various applications (material sciences, finance, etc) involve high-dimensional partial differential equations (PDEs). The typical example in finance is the pricing of a basket option, which can be obtained by solving the Black-Scholes PDE with dimension the number of underlying assets. We propose to investigate an algorithm which has been recently proposed and analyzed in [ACKM06, BLM09] to solve such problems and try to circumvent the curse of dimensionality. The idea is to represent the solution as a sum of tensor products and to compute iteratively the terms of this sum using a greedy algorithm. The resolution of high dimensional partial differential equations is highly related to the representation of high dimensional functions. In Chapter 1, we describe various linear approaches existing in literature to represent high dimensional functions and we introduce the high dimensional problems in finance that we will address in this work.

The method studied in this manuscript is a non-linear approximation method called the Proper Generalized Decomposition. Chapter 2 shows the application of this method to approximate the solution of a linear PDE (the Poisson problem) and also to approximate a square integrable function by a sum of tensor products. A numerical study of this last problem is presented in Chapter 3. The Poisson problem and the approximation of a square integrable function will serve as basis in Chapter 4 for solving the Black-Scholes equation using the PGD approach. In numerical experiments, we obtain results for up to 10 underlyings.

Besides the approximation of the solution to the Black-Scholes equation, we propose a variance reduction method, which permits an important reduction of the variance of the Monte Carlo method for option pricing.

Second theme : Liquidity risk, limit order book modeling and market microstructure

Liquidity risk and market microstructure have become in the past years an important topic in mathematical finance. One possible reason is the deregulation of markets and the competition between them to try to attract as many investors as possible. Thus, quotation rules are changing and, in general, more information is available. In particular, it is possible to know at each time the awaiting orders on some stocks and to have a record of all the past transactions. In this work we study how to use this information to optimally execute buy or sell orders, which is linked to the traders' behaviour that want to minimize their trading cost.

The structure of Limit Order Books (LOB) is very complex. Orders can only be made on a price grid. At each time, the number of waiting buy (or sell) orders for each price is stored. For a given price, orders are executed according to the First In First Out rule, as soon as two orders match together. Thus, since it is really complex, an exhaustive modeling of the LOB dynamics would not lead, for example, to draw conclusions on an optimal trading strategy. One has therefore to propose models that can grasp important features of the LOB structure but that allow to find analytical results.

In [AFS10], Alfonsi, Fruth and Schied have proposed a simple LOB model. In this model, it is possible to explicitly derive the optimal strategy for buying (or selling) a given amount of shares before a given deadline. Basically, one has to split the large buy (or sell) order into smaller ones in order to find the best trade-off between attracting new orders and the price of the orders.

Here, we focus on an extension of the Limit Order Book (LOB) model with general shape introduced by Alfonsi, Fruth and Schied. The additional feature is a time-varying LOB depth that represents a new feature of the LOB highlighted in [JJ88, GM92, HH95, KW96]. We solve the optimal execution problem in this framework for both discrete and continuous time strategies. This gives in particular sufficient conditions to exclude Price Manipulations in the sense of Huberman and Stanzl [HS04] or Transaction-Triggered Price Manipulations (see Alfonsi, Schied and Slynko). These conditions give interesting qualitative insights on how market makers may create price manipulations.

Contents

Part I

Greedy algorithms and application for solving high-dimensional partial differential equations

1

Approximation of high-dimensional functions and the pricing problem

The approximation of high-dimensional functions is an important subject because of the large domain of applications.

The main difficulty for approximating high-dimensional functions is that when the dimension increases, the quantity of information increases exponentially fast with the dimension. This obstacle is known as the curse of dimensionality.

In Section 1.2, we present different approaches proposed in the literature for representing high-dimensional functions. In particular, in Section 1.2, we discuss the linear techniques, the non-linear methods being defined in Chapter 2. We draw your attention on the fact that the non-linear techniques will be the methods used in this manuscript.

Before introducing these linear methods to approximate high-dimensional functions, let us discuss the curse of dimensionality in order to understand the difficulties behind the study of high-dimensional problems.

1.1 The curse of dimensionality

Let us introduce the Hilbert space V. The main idea of the deterministic approaches is to represent solutions $u \in V$ as linear combinations of tensor products. The approximation by a full tensor products writes:

$$u(x_1, x_2 \ldots, x_d) = \sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \ldots \sum_{i_d=1}^{N_d} u^{i_1 i_2 \ldots i_d} \phi_{i_1}(x_1) \phi_{i_2}(x_2) \ldots \phi_{i_d}(x_d) \tag{1.1}$$

where $u^{i_1 i_2 \ldots i_d} \in \mathbb{R}$ for all $i_j = 0, \ldots N_j,\ j = 1, \ldots, d$ and $\left(\phi_{i_j}\right)_{1 \leq i_j \leq N_j}$ are the basis of the vector spaces of dimension N_j for all $j = 1, \ldots, d$ which are fixed. As a consequence, this approach leads to considering a number of degrees of freedom $\mathcal{N}$

$$\mathcal{N} = \prod_{j=1}^{d} N_j \tag{1.2}$$

that grows exponentially in terms of the dimension d.

The following result given by DeVore, Howard and Micchelli [DHM89] allows to see in practice the curse of dimensionality because it shows that a sampling method cannot do better than a certain error estimate depending exponentially on the dimension.

This result is based on the non-linear manifold width. Let X be a normed space and $\mathcal{K} \subset X$ a compact set. Let us consider the maps $E : \mathcal{K} \mapsto \mathbb{R}^N$ for the encoding and $R : \mathbb{R}^N \mapsto X$ for the reconstruction. Introducing the distortion of the pair (E, R) over $\mathcal{K}$

$$\max_{u \in \mathcal{K}} \|u - R(E(u))\|_X,$$

we define the non-linear N-width of $\mathcal{K}$ as

$$d_N(\mathcal{K}) := \inf_{E,R} \max_{u \in \mathcal{K}} \|u - R(E(u))\|_X$$

where the infimum is taken over all the continuous maps (E, R). If $X = L^\infty$ and $\mathcal{K}$ is the unit ball of $C^m([0,1]^d)$, it can be proven that (see [DHM89])

$$cN^{-\frac{m}{d}} \leq d_N(\mathcal{K}) \leq CN^{-\frac{m}{d}}$$

where c and C are two constant that do not depend on N. For a fixed error level, the number of degrees of freedom grows exponentially fast with the dimension. In conclusion, for high-dimensional problems, appropriate approximation tools should be studied.

In this direction, we present in the following Section 1.2 approximation methods that allow to reduce the number of degree of freedom given the tensorial form of their approximated solutions.

1.2 Some approaches to approximate high-dimensional functions

In this section, we present a short survey on methods proposed in the literature for approximating high-dimensional functions. We recall that the approach used in this work to obtain this approximations is presented in Chapter 2.

1.2.1 Sparse grids

The sparse grid method is a numerical discretization technique for multivariate problems. This approach is introduced in [Smo63] and studied by Schwab [PS04] and Zenger [Zen91]. In this part, we present a short survey of this method. See [BG04] for a complete introduction to the sparse grid methods. This approach is also known under the name of hyperbolic cross points or splitting interpolations.

Let us consider $\mathcal{X} = [0,1]$. The use of one-dimensional multilevel (or hierarchical) basis is one of the main ideas of the sparse grid method. In the classical approach, the following standard hat function is employed to construct the hierarchical basis functions

$$\phi(\xi) := \begin{cases} 1 - |\xi|, \text{ if } \xi \in [-1,1] \\ 0, \text{ otherwise.} \end{cases} \tag{1.3}$$

Consequently, we can consider a set of equidistant grids of level m and mesh width $h_m = 2^{-m}$ on $\mathcal{X}$ by introducing the following points:

$$\xi_{m,i} := i2^{-m},\ 0 \le i \le 2^m.$$

Associated to the points $\xi_{m,i}$, we define the basis function $(\phi_{m,i})_{1\le i\le 2^m-1}$ using the standard hat function (1.3)

$$\phi_{m,i}(x) := \phi\left(\frac{x - x_{m,i}}{2^{-m}}\right).$$

We note that this basis is the standard basis of $\mathbb{P}_1$ Lagrange finite element functions with mesh size h_m and with support on $[\xi_{m,i} - h_m, \xi_{m,i} + h_m]$.

Thus, we can define the function spaces

$$V_m := \operatorname{Span}\{\phi_{m,i}, \quad 1 \le i \le 2^m - 1\}$$

and the hierarchical increment spaces W_m

$$W_m := \operatorname{Span}\{\phi_{m,i}, i \in I_m\},$$

where the index set I_m is defined as follows

$$I_m := \{i \in \mathbb{N}, 1 \leq i \leq 2^m - 1, i \text{ odd } \}.$$

Hence, the increment spaces verify the following relation

$$V_m = \bigoplus_{k \leq m} W_k,$$

where the symbol $\oplus$ means that the sum is direct. This decomposition $(\phi_{i,k})_{k \leq m, i \in I_k}$ leads to the hierarchical basis of V_m because any continuous piecewise linear function $u \in V_m$ can be written as

$$u = \sum_{k=1}^{m} \sum_{i \in I_k} u_{k,i} \phi_{k,i},$$

with $u_{k,i} \in \mathbb{R}$ for all $1 \leq k \leq m$ and $i \in I_k$. We remark that the support of all the basis functions $\phi_{k,i}$ spanning W_k are mutually disjoint.

In order to explain, the tensor product construction in high-dimensional spaces, let us introduce $i = (i_1, \ldots, i_d) \in \mathbb{N}^d$ and $k = (k_1, \ldots, k_d) \in \mathbb{N}^d$ two multi-indices. The notation

$$i \leq k$$

means that

$$\forall\, 1 \leq j \leq d,\ i_j \leq k_j$$

Moreover, we will consider the notation $2^i = (2^{i_1}, \ldots, 2^{i_d}) \in \mathbb{N}^d$ and $1 = (1, \ldots, 1) \in \mathbb{N}^p$.

The goal is to construct a multi-dimensional basis on $\mathcal{X}^d = [0,1]^d$ from the one-dimensional hierarchical basis. In order to do that, we consider the d-dimensional tensorization of the one-dimensional basis $(\phi_{k,i})_{1 \leq k \leq m, i \in I_k}$ by introducing $m = (m_1, \ldots, m_d)$ the multi-index denoting the level of discretization in each dimension and the grid points $x_{m,i}$ given by

$$x_{m,i} = (x_{m_1,i_1}, \ldots, x_{m_d,i_d})$$

where $i = (i_1, \ldots, i_d) \in \mathbb{N}^d$ with $1 \leq i \leq 2^m$.

Then, for each grid point $x_{m,i}$, an associated d-dimensional basis function $\phi_{m,i}$ is defined as the product of the one-dimensional basis functions.

$$\phi_{m,i}(x_1,\ldots,x_d) := \prod_{j=1}^{d} \phi_{m_j,i_j}(x_j).$$

Thus, using this basis of functions, we can define the spaces V_m of continuous piecewise d-linear functions

$$V_m := \text{Span}\left\{\phi_{m,i},\ 1 \le i \le 2^m - 1\right\}. \tag{1.4}$$

As in the one-dimensional case, we can define the hierarchical increments W_m as follows

$$W_m := \text{Span}\left\{\phi_{m,i},\ i \in I_m\right\}$$

where $I_m := \left\{i \in \mathbb{N}^d, 1 \le i \le 2^m - 1, i_j \text{ odd for all } 1 \le j \le d\right\}$

Consequently, the spaces V_m verify the property

$$V_m = \bigoplus_{k \le m} W_k,$$

and therefore any function $u \in V_m$ can be written under the form

$$u(x) = \sum_{1 \le k \le m} \sum_{i \in I_k} u_{k,i}\phi_{k,i}(x),\ u_{k,i} \in \mathbb{R}.$$

In order to introduce an optimization with respect to the number of degrees of freedom and the obtained accuracy of the approximation, we consider the sparse grid space $\hat{V}_n$ of level n defined as follows:

$$\hat{V}_n := \bigoplus_{|k|_1 \le n+d-1} W_k.$$

where $|k|_1$ and $|k|_\infty$ are two norms for multi-indices $k = (k_1,\ldots,k_d) \in \mathbb{N}^d$ such that

$$|k|_1 := \sum_{j=1}^{d} |k_j| \text{ and } |k|_\infty := \max_{1 \le j \le d} |k_j|,$$

moreover, in this case, the associated full grid space V_n can be written as

$$V_n := \bigoplus_{|k|_\infty \le n} W_k,$$

The dimension of the space $\hat{V}_n$, that means, the number of degrees of freedom or grid points is given by

$$\begin{aligned}\dim\hat{V}_n &= \sum_{i=0}^{n-1} 2^i \binom{p-1+i}{p-1} \\ &= \mathcal{O}(h_n^{-1}|\log_2(h_n)|^{p-1}).\end{aligned}$$

We remark that the space V_n can be seen as the discretization space associated with a standard $\mathbb{P}_1$ finite element discretization based on a uniform discretization of mesh size $h_n = 2^{-n}$, so the dimension of the space V_n is of the order $\mathcal{O}(h_n^{-p})$. Consequently, the reduction in the number of degrees of freedom is significant by considering $\hat{V}_n$ instead of V_n.

To show the accuracy of the approximation obtained by the sparse grid methods, we introduce the following Sobolev space

$$H^{2,\text{mix}}(\mathcal{X}^d) := \left\{u \in L^2(\mathcal{X}^d), \partial^\alpha u \in L^2(\mathcal{X}^d), \alpha \in \mathbb{N}^d, |\alpha|_\infty \leq 2\right\},$$

and we denote by Π_{V_n} and $\Pi_{\hat{V}_n}$, the $L^2(\mathcal{X}^d)$-orthogonal projector of $L^2(\mathcal{X}^d)$ onto V_n and onto $\hat{V}_n$ respectively.

For all $u \in H^{2,\text{mix}}(\mathcal{X}^d) \cap H_0^1(\mathcal{X}^d)$, the approximation error of the function u on the sparse grid space is

$$\|u - \Pi_{\hat{V}_n} u\|_{L^2(\mathcal{X}^d)} = \mathcal{O}(h_n^2 n^{d-1}), \tag{1.5}$$

and on the full grid space, the accuracy is

$$\|u - \Pi_{V_n} u\|_{L^2(\mathcal{X}^d)} = \mathcal{O}(h_n^2). \tag{1.6}$$

For a given error level, in (1.5), the number of degrees of freedom does not grow exponentially with the dimension. These results (see [BG04]) show the advantage of using the sparse grid space $\hat{V}_n$ with respect to the full grid space V_n because the number of degrees of freedom is strongly reduced while the accuracy is insignificantly deteriorated if the exact solution is regular enough. The efficiency of the sparse tensor products is lost when the solution u is not regular or when the considered mesh is complicated. Consequently, in practice, this method may be difficult to apply for reasons such as the lack of regularity of the solution and the difficulty to implement the associated algorithms.

1.2.2 Canonical, Tucker and Tensor Train decompositions

In this section, we suppose that the Hilbert space V has the following form:

$$V = \bigotimes_{i=1}^{d} V_i. \tag{1.7}$$

i.e., V is tensor product of Hilbert spaces of univariate functions. In other words, for all $1 \leq i \leq d$, the function $u_i \in V_i$ is such that $u_i : x_i \in \Omega_i \mapsto u(x_i)$ and then a function $u \in V$ is such that $u : (x_1, \ldots, x_d) \in \Omega_1 \times \ldots \times \Omega_d \mapsto u(x_1, \ldots, x_d)$. By considering (1.7), the goal is to obtain a number of degrees of freedom which does not depend exponentially on the dimension d.

Canonical decomposition

The canonical decomposition is a classical method to represent in a tensor format a function $u \in V$. This approach looks for a representation of u as follows

$$(x_1, \ldots, x_d) \mapsto u(x_1, \ldots, x_d) = \sum_{k=1}^{r} \left(\bigotimes_{i=1}^{d} u_{i,k} \right) (x_1, \ldots, x_d), \tag{1.8}$$

that is, u is represented by r elementary products of single-variate functions. The number r of products of single-variate functions is called the canonical rank of the function u. In a finite dimensional case with $\dim(V_i) = N$, for all $i = 1, \ldots, d$, we can deduce that the complexity of the decomposition (1.8) is rdN.

Nevertheless, one disadvantage of this approach is that the set of rank-r tensors

$$\mathcal{C}_r := \left\{ u \in V, u(x_1, \ldots, x_d) = \sum_{k=1}^{r} \left(\bigotimes_{i=1}^{d} u_{i,k} \right) (x_1, \ldots, x_d), \forall 1 \leq r, u_{i,k} \in V_i \right\}$$

is not a weakly closed subset of V when $d \geq 3$ and $r \geq 2$, see [dSL08]. Consequently, there may not exists a minimizer to the problem

$$\inf_{\tilde{u} \in \mathcal{C}_r} \|u - \tilde{u}\|_V$$

In the literature, this decomposition is also called CANDECOMP or PARAFAC as in [BK09].

Tucker decomposition

A more robust tensor format decomposition is called the Tucker decomposition that consists in decomposing u as follows:

$$(x_1,\dots,x_d) \mapsto u(x_1,x_2,\dots,x_d) = \sum_{k_1=1}^{r_1} \dots \sum_{k_d=1}^{r_d} c_{k_1,\dots,k_d} \left(\bigotimes_{i=1}^{p} u_{i,k_i} \right) (x_1,\dots,x_p),$$

where for all $1 \leq i \leq d$, $r_i \in \mathbb{N}^*, u_{i,k_i} \in V_i$ for all $1 \leq k_i \leq r_i$ and $c_{k_1,\dots,k_p} \in \mathbb{R}$.

Thus, in the Tucker decomposition, the function u is decomposed over all the possible tensor products between the functions $(u_{i,k_i})_{1 \leq k_i \boldsymbol{r}_i}$ for all $1 \leq i \leq d$.

Let us note $r_T := (r_1,\dots,r_p) \in (\mathbb{N}^*)^p$ the Tucker rank of u. The set of tensors of Tucker rank r_T is weakly closed in V, then the problem of the best approximation has a solution. Nevertheless, we observe that if $r_T = (r,\dots,r)$ and $\dim V_i = N$ for all $1 \leq i \leq d$ this approach leads to a complexity given by $\mathcal{O}(r^d + Nrd)$ that is exponential with the respect to the dimension d. Hence, the Tucker decomposition is not pertinent when d is large.

Tensor train

The following decomposition allows to overcome the exponential complexity obtained in the case of the Tucker decomposition and it is called tensor train decomposition. In this case, the function u is represented as follows:

$$\begin{aligned}(x_1,\dots,x_d) &\mapsto u(x_1,\dots,x_d) \\ &= \sum_{k_1=1}^{r_1} \dots \sum_{k_{d-1}=1}^{r_{d-1}} U_1(x_1,k_1)U_2(k_1,x_2,k_2)\dots U_{d-1}(k_{d-2},x_{d-1},k_{d-1})U_p(k_{d-1},x_d).\end{aligned}$$

The rank of the function u in the case of the tensor train is defined as $r_{TT} = (r_1,\dots,r_{d-1}) \in (\mathbb{N})^{d-1}$. Based on this tensor train decomposition, a function u can be expressed as the following product of matrices

$$u(x_1,\dots,x_d) = U_1(x_1)\dots U_p(x_p)$$

where

$$x_1 \mapsto U_1(x_1) \in \mathbb{R}^{1\times r_1},$$
$$x_2 \mapsto U_2(x_2) \in \mathbb{R}^{r_1\times r_2},$$
$$\dots$$
$$x_{d-1} \mapsto U_{d-1}(x_{d-1})\mathbb{R}^{r_{d-2}\times r_{d-1}},$$
$$x_d \mapsto U_d(x_d) \in \mathbb{R}^{r_{d-1}\times 1}.$$

The set of functions of tensor train rank of at most r_{TT} is a weakly closed subset of V. Moreover, the complexity of this type of decomposition is $\mathcal{O}(r^2Nd)$ if $r_i = r$ and $\dim V_i = N$ for all $i = 1, \dots d$ and not exponential with respect to d as in the case of the Tucker decomposition.

The algorithms used in practice to compute the best approximation of a given tensor in the Tucker or tensor train format, execute successive Singular Value Decomposition problems. We can mention the Higher Order Orthogonal Iteration (HOOI) [LMV00a], the Newton-Grassman approach [ES09] or the Higher Order Singular Value Decomposition (HOSVD) [LMV00b].

1.3 High-dimensional problems in finance

In finance, many of the high-dimensional applications include high-dimensional partial differential equations (PDE). A high-dimensional PDE is an equation that depends on several independent variables. Roughly, the PDEs can be classified in elliptic, parabolic and hyperbolic. In this work, we principally study the parabolic ones, given that in finance the relation between the Black-Scholes model (see Section 1.3.2) and this type of equations give them a strong importance.

The goal of this section is to present some high-dimensional PDEs that appear in finance and that are studied in this manuscript, but before that, it is important to begin by introducing the standard framework used in mathematical finance.

1.3.1 Important concepts in finance

The pricing of financial options is one of the most important problems in financial mathematics. In 1900, Bachelier [Bacal] is the first that shows that for answering this kind of problems it is important to use suitable mathematical techniques. This domain did not have a very strong development until the 70's with the Black-Scholes model developed in 1973 by Merton, Black and Scholes [Mer76, BS73], where they define the price of derivatives as the price needed to hedge them. After that, the financial mathematics domain was driven by the martingale theory developed in the 80's.

One of the most typical examples of derivatives proposed in markets are the options. An European call (resp. put) option is a financial contract that gives to the holder the option and not the obligation to buy (resp. sell) a number of stocks to the price K at time T. The fixed price K is known as the strike and T is the maturity of the option. The underlying is equal to S_T at the maturity T and the option will be exercised if $S_T > K$ (resp. $S_T < K$) and consequently, the holder's gain is given by $S_T - K$ (resp. $K - S_T$) because he can buy (resp. sell) the stock at the price K and after sell (resp. buy) it in the market at the price S_T. Otherwise, the option is not exercised and the gain is equal to 0. Hence, we note that the value of the call at the maturity is given by

$$(S_T - K)_+ = \max(S_T - K, 0).$$

For the counterpart (the bank) which sells the European call option, the goal is to provide the stocks at price K and thus to obtain at the maturity a wealth equal to $(S_T - K)_+$. At the time when the option is sold, the value of the asset S_T is not known and then the question of the pricing, that is, how much the client has to pay to get a call option, is very important.

To answer the question of the pricing, some assumptions are usually considered. The hypothesis that appears in most models is that in a liquid market there is no arbitrage, that means, it is impossible to make profit without taking risks.

Under the assumption of no arbitrage, the price of the call option at time T is given by $(S_T - K)_+$. In general, the price at the maturity of an option is a function of S_T called the payoff. There exist different types of payoff functions,

$$1_{\forall t \in [0,T], S_t \in [a,b]} \phi(S_T), \text{ for barrier options,}$$

$$\phi(S_T, A_T), \text{ where } A_t = \frac{1}{t}\int_0^t S_u du, \text{ for options on the average,}$$

$$\phi(S_T^1, S_T^2, \ldots, S_T^d), \text{ for a basket option,}$$

In this work, we will mainly consider the payoff:

$$\phi(S_T^1, \ldots, S_T^d) = \left(K - \frac{1}{d}\sum_{i=1}^{d} S_T^i\right)_+ \tag{1.9}$$

called a basket put option.

Hence, the function ϕ depends on d different assets which, in general, do not evolve independently.

1.3.2 The Black-Scholes model

In section 1.3.1, we already presented the main ideas of the theory of options in the framework of financial mathematics. The goal of this section is to develop the relationship between theory and mathematical equations that are obtained based on the Black-Scholes model.

A partial differential equation for the option pricing problem

In this section, we present the framework that allows to find a PDE for pricing European options. In order to do that, we introduce the classical Black-Scholes model that takes into account a risky asset with price S_t at the time t and a risk-free asset which price at time t is S_t^0. In this model, S_t and S_t^0 are such that:

$$dS_t = \mu S_t dt + \sigma S_t dB_t,$$

and

$$dS_t^0 = r S_t^0 dt$$

where B_t is a Brownian motion in a probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and μ (the mean rate of return), $\sigma > 0$ (the volatility) and r (the risk-free interest rate) are three constants. This framework can be generalized to the case where r, μ and σ are functions of the time t and the stock S_t under suitable smoothness assumptions. The filtration $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration of the Brownian motion B_t.

Let us introduce the risk-free probability measure $\mathbb{Q}$ defined by its Radom-Nikodim derivative with respect to $\mathbb{P}$ as follows

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = \exp\left(\int_0^t \frac{\mu - r}{\sigma} dB_s + \int_0^t \left(\frac{\mu - r}{\sigma}\right)^2 ds\right) \tag{1.10}$$

This new probability $\mathbb{Q}$ is one of the key tools to obtain the results of the Black-Scholes model.

If we introduce the stochastic process $W_t = B_t + \frac{\mu - r}{\sigma} t$, the process S_t satisfies the following stochastic differential equation under the probability $\mathbb{Q}$

$$dS_t = S_t(r dt + \sigma dW_t) \tag{1.11}$$

where W_t is a Brownian motion and $\frac{S_t}{S_t^0}$ is a martingale under the probability $\mathbb{Q}$.

Since we are interested in the pricing of an option, the Black-Scholes model is studied in the interval $[0, T]$ where T is the maturity of the option. Thus, the solution of the equation (1.11) is given by

$$S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}, \quad t \in [0, T]. \tag{1.12}$$

where S_0 is the value of the asset at time $t = 0$.

Specifically, we can observe that the process $(S_t)_{t\geq 0}$ satisfies the equation (1.11) if the process $(\log(S_t))_{t\geq 0}$ is a Brownian motion, not necessarily standard. Looking at the expression of the process S_t obtained in (1.12), we find the assumptions of the Black-Scholes model on the evolution of the asset. These are:

- Continuity of the trajectories.
- If $u \leq t$, $\frac{S_t}{S_u}$ is independent of the sigma-algebra $\sigma\left(S_v, v \leq u\right)$
- If $u \leq t$, $\frac{(S_t - S_u)}{S_u}$ and $\frac{(S_{t-u} - S_0)}{S_0}$ have the same law.

Let us define now a strategy that allows to generate a portfolio. A strategy is a process $(\mathcal{H}_t)_{0\leq t\leq T} = (H_t, H_t^0) \in \mathbb{R}^2$ adapted to the natural filtration of the Brownian motion, with H_t risky assets and H_t^0 non-risky assets. Hence, the value of the associated portfolio P_t is

$$P_t = H_t S_t + H_t^0 S_t^0 \tag{1.13}$$

The considered portfolio P_t is assumed to be self-financing, that means, any change on this portfolio is done with no exogenous supply or withdrawal of money. Mathematically, this assumption writes

$$dP_t = H_t dS_t + H_t^0 dS_t^0, \tag{1.14}$$

and using this equation (1.14) it is possible to show that $\frac{P_t}{S_t^0}$ is a martingale.

As we mentioned in section 1.3.1, the great idea of the Black-Scholes model is to put together the pricing of the option and the quantity of money needed to hedge it. Theoretically, for a given function ϕ (the payoff) and a given maturity T, it is possible to create a self-financing portfolio such that $P_T = \phi(S_T)$. This result is obtained using the martingale representation theorem, the fact that ϕ is $\mathcal{F}_T$-measurable and that $\frac{P_t}{S_t^0}$ is a martingale. Using the martingale property, the value of the portfolio at the time t is

$$P_t = \mathbb{E}\left[e^{-(T-t)r}\phi(S_T)|\mathcal{F}_t\right] \tag{1.15}$$

Using the so-called arbitrage-free principle, it is possible to show that P_t has to be the price at time t of the option which allows the holder to obtain the payoff $\phi(S_T)$ at the maturity T.

A PDE formulation of the pricing option problem can be obtained because of the Markov property of the process S_t. This Markovianity means that the expectation of any function of $(S_t)_{0\leq t\leq T}$ conditionally to $\mathcal{F}_t$ is a function of the price of the asset S_t at time t. In other words, this implies that

$$P_t = p(t, S_t) \tag{1.16}$$

where p is a function of $t \in [0, T]$ and $S \in [0, \infty)$, known as the pricing function of the option. We remark that the pricing function p is a deterministic function defined for all values of $t \geq 0$ and $S \geq 0$.

As a consequence of the Markov property of S_t, we can re-write the pricing function p under the form:

$$p(t, x) = \mathbb{E}\left[e^{r(T-t)}\phi(S_T^{t,x})\right]$$

where the process $S_u^{t,x}$ is the solution of the equation (1.11) starting from x at time t, or equivalently,

$$\begin{cases} dS_u^{t,x} = S_u^{t,x}(rdu + \sigma dW_u), \quad u \geq t, \\ S_t^{t,x} = x \end{cases}$$

Using Itô's calculus and the fact that $\frac{P_t}{S_t^0}$ is a martingale, we deduce that p has to verify the following PDE:

$$\begin{cases} \frac{\partial p}{\partial t} + rS\frac{\partial p}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 p}{\partial S^2} - rp = 0, \\ p(T, S) = \phi(S). \end{cases} \tag{1.17}$$

It is possible to show that if p verifies (1.17), then $p(t, S_t)$ is the value of a self-financing portfolio such that its value is equal to $\phi(S_T)$ at time T.

The Black-Scholes formula

In this section, we present the formulas obtained for the pricing of European options. The fact that relatively simple expressions give the price of these financial contracts is one of the most important features of this model.

Let us note $P(t,S)$ the price of the option with maturity T of payoff ϕ. Assuming that r and $\sigma > 0$ are constants, the price of the option in the Black-Scholes equation is given by

$$P(t,S) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[\phi\left(Se^{r(T-t)}e^{\sigma(W_T-W_t)-\frac{\sigma^2}{2}(T-t)}\right)\right] \tag{1.18}$$

The expression (1.18) can be written under the following form

$$P(t,S) = \frac{1}{\sqrt{2\pi}}e^{-r(T-t)}\int_{\mathbb{R}}\phi\left(Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}y} - Ke^{-r(T-t)}\right)e^{-\frac{y^2}{2}}dy$$

because under the probability $\mathbb{Q}$, $W_T - W_t$ is a centered Gaussian random variable with variance $T-t$.

If we take the case of a put option, that means that the payoff ϕ is such that $\phi(S_t) = (S_t - K)_+$, we can get that

$$C(t,S) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2}\left(Se^{-\frac{\sigma^2}{2}(T-t)-\sigma y\sqrt{T-t}} - Ke^{-r(T-t)}\right)e^{-\frac{y^2}{2}}dy, \tag{1.19}$$

where

$$d_1 = \frac{\log\left(\frac{S}{K}\right) + (r+\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \text{ and } d_2 = d_1 - \sigma\sqrt{T-t} \tag{1.20}$$

The Black-Scholes formula can be derived from (1.19).

Proposition 1.1. *If we assume that $\sigma > 0$ and r are two constants, the price of a European call option is given by*

$$C(t,S) = S\mathcal{N}(d_1) - Ke^{-r(T-t)}\mathcal{N}(d_2), \tag{1.21}$$

and for the case of a European put option, the price is

$$P(t,S) = -S\mathcal{N}(-d_1) + Ke^{-r(T-t)}\mathcal{N}(-d_2). \tag{1.22}$$

where d_1 and d_2 are defined by (1.20) *and $\mathcal{N}$ is the cumulative distribution function of a centered Gaussian distribution with variance equal to* 1.

Finally, we can remark that if r and σ are functions of time, the formulas (1.21) and (1.22) are still valid, replacing the expressions (1.20) by

$$d_1 = \frac{\log\left(\frac{S}{K}\right) + \int_t^T r_u du + \frac{1}{2}\int_t^T \sigma_u^2 du}{\sqrt{\int_t^T \sigma_u^2 du}} \text{ and } d_2 = d_1 - \sqrt{\int_t^T \sigma_u^2 du} \tag{1.23}$$

1.3.3 High-dimensional partial differential equations in finance

The goal of this section is to extend the model presented in the previous Section 1.3.2 to the case when d-assets are considered as the underlyings of an option.

Model with d risky assets

We consider that there exist d risky assets whose prices at time t are written $S_i(t)$. We assume that for all i such that $1 \leq i \leq d$, the price $S_i(t)$ verifies the following stochastic differential equation:

$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t) dW_i(t) \tag{1.24}$$

We have to point out that $(W_i(t))$ for $1 \leq i \leq d$ are correlated Brownian motions defined on a probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Let us introduce ρ_{ij} the correlation factor between $W_i(t)$ and $W_j(t)$. That is

$$\rho_{ij} = \frac{\mathbb{E}\left[W_i(t)W_j(t)\right]}{t}$$

We remark that $-1 \leq \rho_{ij} \leq 1$ and $\rho_{ii} = 1$. In addition, the volatilities σ_i, for $1 \leq i \leq d$ are positive constants.

As in Section 1.3.2, we are interested in the study of an European option, but on d underlyings, with $d > 1$. If T is the maturity of the option of payoff $f(S_1(T), \ldots, S_d(T))$, it is possible to find a risk-free probability measure $\mathbb{Q}$ that allows to write the price of the option at time t as follows:

$$P_t = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)} f(S_1(T), \ldots, S_d(T)) | \mathcal{F}_t\right] \tag{1.25}$$

Let us find, as in the case presented in Section 1.3.2, the linear parabolic partial differential equation on $d+1$ variables linked with the equation (1.25) and that allows to find the price P_t at time t of the option. In order to do that, we cite some results given in [AP05].

Proposition 1.2. *Let us introduce the differential operator*

$$L : f \mapsto \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \Xi_{ij} S_i S_j \frac{\partial^2 f}{\partial S_i S_j} + r\sum_{i=1}^{d} S_i \frac{\partial f}{\partial S_i}, \tag{1.26}$$

where $\Xi_{ij} = \rho_{ij}\sigma_i\sigma_j$. Then, for all function $u : (S_1, \ldots, S_d, t) \mapsto u(S_1, \ldots, S_d, t), u \in C^{2,1}(\mathbb{R}_+^d \times [0,T))$ verifying $|S_i \frac{\partial u}{\partial S_i}| \leq C(1 + |S_i|)$, $i = 1, \ldots, d$ with C that does not depend on t, the process

$$M_t = e^{-rt}u(S_1(t), \ldots, S_d(t)) - \int_0^t e^{-r\tau}\left(\frac{\partial u}{\partial t} + Lu - ru\right)(S_1(t), \ldots, S_d(t))\tau$$

is a martingale under the filtration $\mathcal{F}_t$.

Theorem 1.1. *Let P be a continuous function, $P \in C^{2,1}(\mathbb{R}_+^d \times [0,T))$ such that $|S_i \frac{\partial P}{\partial S_i}| \leq C(1 + S_i)$ where C does not depend on t. Assuming that P verifies the equation*

$$\left(\frac{\partial P}{\partial t} + LP - rP\right)(S_1, \ldots, S_d, t) = 0, \ \ t < T, (S_1, \ldots, S_d) \in \mathbb{R}_+^d \tag{1.27}$$

and satisfies the Cauchy condition

$$P(S_1, \ldots, S_d, T) = f(S_1, \ldots, S_d), \ \ (S_1, \ldots, S_d) \in \mathbb{R}_+^d,$$

then the price of the European option given by (1.25) verifies

$$P_t = P(S_1(t), \ldots, S_d(t), t)$$

It should be outlined that in the case of a basket option, analytical formulas such as those presented for the case of one asset can no longer be found. This implies the use of numerical methods in order to approximate the price.

Free boundary problems

In some applications we do not only need to find the solution of a PDE but it is also necessary to define constraints on an unknown boundary. The problem of the execution of an American option can be studied with this type of problem.

It is possible to show (see [AP05]) that the partial differential equation for the pricing of the American option can be expressed under the following form:

$$\begin{cases} \min\left(Lu - \dfrac{\partial u}{\partial t}, \phi - u\right) = 0, \quad \text{in } \mathcal{R}_T := (0,T) \times \mathbb{R}^d \\ u(0,x) = \phi(x), \ \ x \in \mathbb{R}^d \end{cases} \tag{1.28}$$

where the parabolic operator L is defined by (1.26).

From equation (1.28), we deduce that $u \geq \phi$ and then the region $\mathcal{R}_T$ is divided in two parts: the so-called exercise region where $u = \phi$ and the continuation region where $u > \phi$ and $Lu - \frac{\partial u}{\partial t} = 0$. We remark that, in the continuation region, the price of the option verifies the Black-Scholes PDE.

In fact, the problem (1.28) is equivalent to

$$\begin{cases} Lu - \dfrac{\partial u}{\partial t} \leq 0, \quad \text{in } \mathcal{R}_T \\ u \geq \phi, \quad \text{in } \mathcal{R}_T \\ (u - \phi)\left(Lu - \dfrac{\partial u}{\partial t}\right) = 0, \quad \text{in } \mathcal{R}_T \\ u(0, x) = \phi(0, x), \quad x \in \mathbb{R}^d \end{cases} \tag{1.29}$$

This type of problem is known as the obstacle problem and is presented in Section 5.2 where we propose an application of the algorithm introduced in Chapter 2 for treating the problem of the American options.

2

A nonlinear approximation method for solving high-dimensional partial differential equations

Many problems of interest for various applications such as kinetic models, molecular dynamics, quantum mechanics, uncertainty quantification and finance involve high-dimensional partial differential equations.

It is well known that when the number of variables of a PDEs is very large, standard algorithms such as finite differences and finite elements cannot be used in practice to solve them. As we discussed in Section 1.1, the reason is the curse of dimensionality, in other words, the number of unknowns typically grows exponentially with respect to the problem's dimension and rapidly exceeds the limited storage capacity.

The main goal of this Section is to present an algorithm which has been recently proposed by Chinesta *et al.* [ACKM06] for solving high-dimensional Fokker-Planck equations in the context of kinetic models for polymers and by Nouy *et al.* [Nou10] in uncertainty quantification framework based on previous works by Ladevèze [Lad99]. This approach is also studied in [BLM09] to try to circumvent the curse of dimensionality for the Poisson problem in high-dimension. This approach is a nonlinear approximation method that we will call below the *Proper Generalized Decomposition* (PGD). It is related to the so-called greedy algorithms introduced in nonlinear approximation theory, see for example [Tem08]. The main idea of the PGD is to represent the solution as a sum of tensor products:

$$\begin{aligned} u(x_1,\dots,x_d) &= \sum_{k\geq 1} r_k^1(x_1) r_k^2(x_2)\dots r_k^d(x_d) \\ &= \sum_{k\geq 1} \left(r_k^1 \otimes r_k^2 \otimes \dots \otimes r_k^d \right)(x_1,\dots,x_d) \end{aligned} \tag{2.1}$$

and to compute iteratively each term of this sum using a greedy algorithm. The algorithm of the PGD method can be applied to any PDE which admits a variational interpretation as a minimization problem. The practical interest of this algorithm has been demonstrated in various contexts (see for example [ACKM06] for applications in fluid mechanics).

One contribution of this work is to complete the first application of this algorithm in finance, investigating the interest of this approach for option pricing. Our aim is to study the problem of pricing vanilla basket options of European type by solving the Black-Scholes equation and to propose in addition a variance reduction method for the pricing of the same type of financial products.

This application in finance leads us to consider two extensions of the standard algorithm, which apply to symmetric linear partial differential equations: (i) non-symmetric linear problems to value European options, (ii) nonlinear variational problems to price American options.

In this chapter, our goal is to present this nonlinear approximation method that we study in this work from a general mathematical point of view.

In what follows, we introduce the approach that we study to solve high-dimensional PDEs that arise in finance. This method is called Proper Generalized Decomposition and it is connected to the greedy algorithms proposed in the nonlinear approximation theory by Temlyakov in [Tem08], Davis *et al.* in [ADM97] and Barron *et al.* in [BCDD08]. Other related works are the methods based on looking for the best n-term approximation of operators as in [BK09] and [BM02].

We begin by presenting the greedy algorithms in a general framework and the link with the PGD. After that, we define the PGD and we apply it, as an example, to the Poisson problem. Finally, we recall results on convergence and speed of convergence for the Proper Generalized Decomposition.

2.1 Greedy algorithms

In what follows, we present an introduction to greedy algorithms. See [Tem08], [DT96] or [BCDD08] for more details. We consider V a real Hilbert space associated with the inner product $\langle ., . \rangle_V$ and with the norm $\|.\|_V$. We define a dictionary $\mathcal{D}$ as a family of functions from V such that all the elements of the dictionary $\mathcal{D}$ are normalized, that means $\|g\|_V = 1$ for all $g \in \mathcal{D}$ and $\mathrm{Span}(\mathcal{D}) = V$. We also assume that the dictionary is symmetric namely that $g \in \mathcal{D} \Rightarrow -g \in \mathcal{D}$.

The problem tackled by the greedy algorithms theory is the problem of approximating a function $u \in V$ by a finite linear combination of elements of the dictionary $\mathcal{D}$. Thus, to analyze the greedy algorithms, we introduce the best n-term approximation u_n of the function $u \in V$ where u_n is a linear combination of at most n elements of the dictionary $\mathcal{D}$. Mathematically, it amounts to looking for the functions $g_1, \ldots, g_n \in \mathcal{D}$ such that they minimize the following error:

$$(g_1, \ldots, g_n) \in \underset{(d_1,\ldots,d_n)\in\mathcal{D}}{\operatorname{argmin}} \|u - P_{d_1,\ldots,d_n} u\|_V$$

where $P_{d_1,\dots,d_n}$ is the orthogonal projector on $\mathrm{Span}\{d_1,\dots,d_n\}$ with respect to the inner product of V.

Instead of fixing the functions $(d_1,\dots,d_n) \in \mathcal{D}$ as in the linear case where the best approximation is the projection $P_{d_1,\dots,d_n}$, the nonlinear framework lets the functions $(d_1,\dots,d_n) \in \mathcal{D}$ depend on the function u that has to be approximated. Hence, the principle of the greedy algorithms is to look iteratively for the best element in the dictionary and, in this way, they propose a constructive way to find the solution for the problem of approximating the function $u \in V$.

In the sequel, we assume that for any function $u \in V$, there exists an element $g \in \mathcal{D}$ such that

$$g \in \underset{d\in\mathcal{D}}{\mathrm{argmax}} \langle u, d\rangle_V \tag{2.2}$$

Thus, if g verifies (2.2), we can deduce directly that

$$(g, \langle u, g\rangle_V) \in \underset{(d,\lambda)\in\mathcal{D}\times\mathbb{R}}{\mathrm{argmin}} \|u - \lambda d\|_V.$$

There are several versions of these greedy algorithms which are introduced in [DT96]. Here, we present two classical versions: the *Pure Greedy Algorithm* (PGA) and the *Orthogonal Greedy Algorithm* (OGA).

Pure greedy algorithm (PGA):

1. Set $r_0^p := u$, $u_0^p := 0$ and $n = 1$. Choose $\epsilon > 0$.
2. Find $g_n^p \in \mathcal{D}$ such that
$$g_n^p \in \underset{g\in\mathcal{D}}{\mathrm{argmax}} \langle r_{n-1}^p, g\rangle_V$$
3. Define $u_n^p := u_{n-1}^p + \langle r_{n-1}^p, g_n^p\rangle_V g_n^p$ and $r_n^p := r_{n-1}^p - \langle r_{n-1}^p, g_n^p\rangle_V g_n^p$.
4. If $\|r_n^p\| \leq \epsilon \|u_n^p\|$, then stop. Otherwise, $n = n + 1$ and return to step 2.

Orthogonal greedy algorithm (OGA):

1. Set $r_0^o := u$, $u_0^o := 0$ and $n = 1$. Choose $\epsilon > 0$.
2. Find $g_n^o \in \mathcal{D}$ such that
$$g_n^o \in \underset{g\in\mathcal{D}}{\mathrm{argmax}} \langle r_{n-1}^o, g\rangle.$$
3. Define $H_n^o := \mathrm{Span}\{g_i^o, 1 \leq i \leq n\}$, $u_n^o := P_{H_n^o}(u)$ and $r_n^o := u - P_{H_n^o}(u)$.
4. If $\|r_n^o\| \leq \epsilon \|u_n^o\|$, then stop. Otherwise, $n = n + 1$ and return to step 2.

The above algorithms are greedy since the basis of vectors used to approximate g is built incrementally; at each iteration a new vector is added, but former vectors are never removed nor modified. We remark that for a general dictionary $\mathcal{D}$ ($\mathcal{D}$ is not an orthonormal basis), the solution obtained after n iterations of the algorithm is generally not the best rank-n approximation (2.14).

The difference between the OGA and the PGA is that the OGA takes the Galerkin projection on the functions $g_1^o, \dots, g_n^o$ generated at each iteration. The first and second steps are the same for the OGA and the PGA.

The convergence of these algorithms is proved for all $u \in V$ as we can see in the following theorem.

Theorem 2.1. *For any dictionary $\mathcal{D}$ and any $u \in V$, we have that*

$$\textit{PGA: } \|r_n^p\| = \|u - u_n^p\| \underset{n\to\infty}{\longrightarrow} 0,$$
$$\textit{OGA: } \|r_n^o\| = \|u - u_n^o\| \underset{n\to\infty}{\longrightarrow} 0,$$

It is also possible to prove convergence rates for these algorithms. In order to do that, it is necessary to define a functional space for the function u, adapted to the convergence analysis. For a general dictionary $\mathcal{D}$, we introduce the following class of functions

$$\mathcal{A}_0^1(\mathcal{D}, M) := \left\{ u \in V : u = \sum_{k\in\Lambda} c_k v_k, v_k \in \mathcal{D}, \#\Lambda < \infty \text{ and } \sum_{k\in\Lambda} |c_k| \leq M \right\}$$

Let us also introduce the space $\mathcal{A}^1(\mathcal{D}, M)$ as the closure in H of $\mathcal{A}_0^1(\mathcal{D}, M)$, that is, $\mathcal{A}^1(\mathcal{D}) := \cup_{M>0}\mathcal{A}^1(\mathcal{D}, M)$, or, equivalently,

$$\mathcal{A}^1(\mathcal{D}) = \left\{ u \in V, u = \sum_{k=1}^{\infty} r_k^1 \otimes r_k^2 \otimes \dots \otimes r_k^{(d)}, \sum_{k=1}^{\infty} \|r_k^1 \otimes r_k^2 \otimes \dots \otimes r_k^{(d)}\|_V < \infty \right\},$$

Thus, the following result proved in [DT96] holds:

Theorem 2.2. *For a general dictionary $\mathcal{D}$ in V, the following estimates can be deduced: For a function $u \in \mathcal{A}^1$, there exists a constant $M > 0$ such that*

$$\|r_n^p\| = \|u - u_n^p\| \leq M n^{-1/6},$$
$$\|r_n^o\| = \|u - u_n^o\| \leq M n^{-1/2},$$

for all $n \in \mathbb{N}^*$.

We note that the constant M depends on the norm $\|u\|_{\mathcal{A}^1}$. In [KT99], Konyagin and Temlyakov obtain a better estimate for the PGA

$$\|r_n^p\| = \|u - u_n^p\| \leq M n^{-11/62}.$$

2.2 The Proper Generalized Decomposition

As we said in the introduction of this Section, this approach has been recently proposed by Chinesta *et al.* in [ACKM06] to solve high-dimensional Fokker-Planck equations in the context of kinetic models for polymers and Nouy in [Nou10] in the context of uncertainty quantification following previous works by Ladevèze [Lad99].

In general, let V be a Hilbert space of multivariate functions $u(x_1, \dots, x_d)$ and let $V_1, \dots, V_d$ be Hilbert spaces of single-variate functions depending on the one-dimensional variable x_i. One of the principles of the PGD is to choose the dictionary of functions as the set of tensor products

$$\mathcal{D} := \left\{ r^1 \otimes \dots \otimes r^d | r^1 \in V_1, \dots r^d \in V_d, \|r^1 \otimes \dots \otimes r^d\|_V = 1 \right\} \tag{2.3}$$

where $r^1 \otimes r^2 \otimes \dots \otimes r^d(x_1, x_2 \dots, x_d) = r^1(x_1) r^2(x_2) \dots r^d(x_d)$.

Let us define the following set of simple products

$$\Sigma := \left\{ r^1 \otimes \dots \otimes r^d, r^1 \in V_1, \dots, r^d \in V_d \right\} \tag{2.4}$$

Under the assumptions

(A1) $\Sigma \subset V$,

(A2) $\overline{\text{Span}\Sigma}^{\|\cdot\|_V} = V$,

(A3) for all sequences of Σ bounded in V, there exists a subsequence which weakly converges in V towards an element of Σ,

the dictionary $\mathcal{D}$ defined in (2.3) is a well-defined dictionary of V. We note that the assumption (A3) implies that the problems of the type (2.2) have at least one solution.

The PGD is based on two main ideas. The first one is to extend the solution as a sum tensor products of lower-dimensional functions

$$u_n(x_1, x_2, \ldots, x_d) = \sum_{k=1}^{n} r_k^1 \otimes r_k^2 \ldots \otimes r_k^d(x_1, x_2, \ldots, x_d) \tag{2.5}$$

where for all $i = 1, \ldots, d$ and $k = 1, \ldots, n$, the functions $r_k^{(i)} \in V_i$. Consequently, the function u_n is a separated representation of the solution $u \in V$.

The second idea is to recast the original problem (in our case a high-dimensional PDE) as a minimization problem:

$$u = \operatorname*{argmin}_{v \in V} \mathcal{E}(v), \tag{2.6}$$

where $\mathcal{E} : V \mapsto \mathbb{R}$ is functional with a unique global minimizer $u \in V$.

In general, to compute u_n in the separated form (2.5), u_n being the approximation of u solution of the problem (2.6), we propose to use the PGD that is defined as follows:

Iterate on $n \geq 1$

$$(r_n^1, r_n^2, \ldots, r_n^d) \in \operatorname*{argmin}_{r^1 \in V_1, r^2 \in V_2, \ldots, r^d \in V_d} \mathcal{E}\left(\sum_{k=1}^{n-1} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d + r^1 \otimes r^2 \otimes \ldots r^d\right), \tag{2.7}$$

where V, $V_1, \ldots, V_d$ are Hilbert spaces.

The principle of the algorithm is to look iteratively for the best tensor product and this leads to a nonlinear approximation method which gives the solution u_n as defined in (2.5). We remark that the implementation of (2.7) amounts to applying the Pure Greedy Algorithm in the Hilbert space V and for the dictionary $\mathcal{D}$ defined by (2.3).

Instead of solving the minimization problem (2.7), we solve the first-order optimality conditions of this minimization problem, namely the Euler equation. This yields to a system of equations where the number of degrees of freedom does not grow exponentially with respect to the dimension. This fact will be very important in order to reach high-dimensional frameworks in practical applications. More precisely, the Euler equation writes as a system of d nonlinear equations, where d is the dimension considered. The maximum dimension that can be treated by this technique is limited by the fact that a system of d nonlinear equations has to be solved. We also note that the solutions of the Euler equation are not necessarily solutions of the minimization problem given the nonlinearity of the tensor product space $V_1 \otimes \ldots \otimes V_d$.

Thus, the greedy algorithm can be stated as follows: For $n \geq 0$, find $\left(r_n^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d\right) \in H_1^0(\Omega_1) \times H_1^0(\Omega_2) \times \ldots \times H_1^0(\Omega_d)$ such that

$$(r_n^1, r_n^2, \ldots, r_n^d) \in \underset{r^1\in V_1,\, r^2\in V_2,\ldots,r^d\in V_d}{\operatorname{argmin}} \mathcal{E}\left(\sum_{k=1}^{n-1} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d + r^1 \otimes r^2 \otimes \ldots \otimes r^d\right), \tag{2.8}$$

Remark 2.1. *Curse of dimensionality*

As we already said, in the PGD approach we look for a solution under the form:

$$u_n(x_1, x_2, \ldots, x_d) = \sum_{k=1}^{n} r_k^1 \otimes r_k^2 \ldots \otimes r_k^d(x_1, \ldots, x_d) \tag{2.9}$$

In practice, the functions r_k^i are obtained as linear combinations of the basis functions $(\phi_{j_i}^l)_{1\le l\le N_i}$. Thus, we introduce the space $V_{x_i}^{h_i}$ as the finite element spaces used to discretize the Hilbert spaces V_{x_i} that are spaces of functions depending on the one-dimensional variable x_i

$$V_{x_i}^{h_i} = Span\{\phi_j, 0 \le j \le N_i\}$$

where h_i is the parameter of discretization $h_i = \frac{1}{N_i}$.

At the end, the problem of computing the approximation (2.9) *leads to solving a problem of dimension $\tilde{\mathcal{N}}$ such that*

$$\tilde{\mathcal{N}} = n\sum_{i=1}^{d} N_i \tag{2.10}$$

which remains lower compared with $\mathcal{N}$ defined in (1.2)*, if n is small enough.*

2.3 Some particular cases: the Singular Value Decomposition and the general linear case

2.3.1 Tensor product of spaces

Let us define the inner product $\langle . \, , \, . \rangle_\otimes$ associated to the space $\mathrm{Span}(\Sigma)$ as follows:

$$\forall \left(r^1, r^2, \ldots, r^{(d)}\right), \left(\tilde{r}^1, \tilde{r}^2, \ldots, \tilde{r}^d\right) \in V_1 \times V_2 \times \ldots \times V_d,$$
$$\langle r^1 \otimes r^2 \otimes \ldots \otimes r^d, \tilde{r}^1 \otimes \tilde{r}^2 \otimes \ldots \otimes \tilde{r}^d\rangle_\otimes = \langle r^1, \tilde{r}^1\rangle_{V_1} \langle r^2, \tilde{r}^2\rangle_{V_2} \ldots \langle r^d, \tilde{r}^d\rangle_{V_d}$$

Likewise, we define the related norm $\|\|_\otimes$ called cross-norm as

$$\forall \left(r^1, r^2, \ldots, r^{(d)}\right) \in V_1 \times V_2 \times \ldots \times V_d,\ \|r^1, r^2, \ldots, r^{(d)}\|_{\otimes} = \|r^1\|_{V_1}\|r^2\|_{V_2} \ldots \|r^{(d)}\|_{V_d}. \tag{2.11}$$

Thus, the tensor product of the spaces $V_1, V_2, \ldots, V_d$, noted by $V_1 \otimes V_2 \otimes \ldots \otimes V_d$ and defined as $\left(\overline{\mathrm{Span}(\Sigma)}^{\|\cdot\|_{\otimes}}, \langle \cdot, \cdot \rangle_{\otimes}\right)$ is a Hilbert space.

In this section, we discuss two cases in which the greedy algorithm satisfies important properties.

2.3.2 The Singular Value Decomposition case

In this case we consider that

$$\mathcal{E}(v) = \|u - v\|_V^2 \tag{2.12}$$

where $V = V_1 \otimes V_2$, that means, the product of only two Hilbert spaces. Moreover, we assume that the norm $\|\cdot\|_V$ is a cross-norm following the definition (2.11) in Section 2.3.1.

In this case, the pairs $(r_n^1, r_n^2) \in V_1 \times V_2$, defined by (2.7), verify the following orthogonality relation:

$$\langle r_n^1, r_m^1 \rangle_{V_1} = \langle r_n^2, r_m^2 \rangle_{V_2} = 0, \forall n \neq m \tag{2.13}$$

This orthogonality property has several consequences:

– The Pure Greedy Algorithm and the Orthogonal Greedy Algorithm are equivalent.
– The decomposition of the function u as follows

$$u = \sum_{k=1}^{\infty} r_k^1 \otimes r_k^2$$

is unique.
– At iteration n, the approximation $u_n = \sum_{k=1}^{n} r_k^1 \otimes r_k^2$ is the best rank-n term approximation of u

$$\|u - \sum_{k=1}^{n} r_k^1 \otimes r_k^2\|_V = \inf_{(\tilde{r}_k^1, \tilde{r}_k^2) \in V_1 \times V_2,\ 1 \leq k \leq n} \|u - \sum_{k=1}^{n} \tilde{r}_k^1 \otimes \tilde{r}_k^2\|_V. \tag{2.14}$$

It is also possible to deduce that

– The solutions to the Euler-Lagrange equation which maximize the L^2-norm $\left(\int_\Omega |r \otimes s|^2\right)^{1/2}$ are exactly the solutions to the minimization problem.
– In dimension $d = 2$, the solutions to the Euler-Lagrange equation that satisfy the second optimality conditions are the solutions of the minimization problem.

– Let $\lambda_k = \|r_k^1 \otimes r_k^2\|_V$ for all $1 \leq k \leq \infty$. The sequence $(\lambda_k)_k \in \mathbb{N}^*$ is non-increasing and the convergence rate of the algorithm is related to this sequence as is showed in (2.15):

$$\|u - u_N\|_V^2 = \sum_{k=n+1}^{\infty} \|r_k^1 \otimes r_k^2\|^2 = \sum_{k=n+1}^{\infty} \lambda_k^2. \tag{2.15}$$

2.3.3 The linear case

In this case, we consider again a quadratic functional $\mathcal{E}(v) = \|u - v\|_V^2$ but V is the product of more than two Hilbert spaces or the norm is not a cross-norm. Here, the convergence results (2.1) and (2.2) hold but the orthogonality property (2.13) is no longer verified. This implies that the PGA and the OGA are not equivalent. The greedy algorithms do not give as solution the best rank-n decomposition as is defined in (2.14) and the sequence $\lambda_k = \|r_k^1 \otimes r_k^2\|_V$ for all $1 \leq k \leq \infty$ is not necessarily non-increasing.

2.4 Other cases of application for the Proper Generalized Decomposition

The work of Cancès, Ehrlacher and Lelièvre [CEL12] study the case when the functional $\mathcal{E}$ is not supposed to be a quadratic energy functional. They extend the work [BLM09] by considering a general strongly convex energy functional. In this section, we give an outline of this extension. In particular, these results can be used to solve an obstacle problem with uncertainty with a large number of random parameters. As the pricing of American options can be written as an obstacle problem (see Section 5.2), the method proposed in [CEL12] can also be applied for the pricing of American options.

Let us consider the following assumptions:

(A4) The energy functional $\mathcal{E}$ is differentiable and strongly convex. Mathematically, that means that there exists $\alpha > 0$ such that

$$\forall v, w \in V, \mathcal{E}(v) \leq \mathcal{E}(u) + \langle \nabla \mathcal{E}(w), v - w \rangle_V + \frac{\alpha}{2} \|v - w\|_V^2.$$

(A5) The gradient of $\mathcal{E}$ is Lipschitz on bounded sets: for each bounded subset $K \subset V$, there exists a constant L_K such that

$$\forall v, w \in K, \|\nabla \mathcal{E}(v) - \nabla \mathcal{E}(w)\|_V \leq L_K \|v - w\|_V.$$

Thus, Cancès, Ehrlacher and Lelièvre prove in [CEL12] the following theorem:

Theorem 2.3. *If the conditions (A1)-(A5) are verified, then the iterations of the algorithm* (2.7) *are well-defined, in the sense that* (2.7) *has at least one minimizer* $(r_n^1, r_n^2, \ldots, r_n^d)$. *Moreover, the sequence* $(u_n)_{n\in\mathbb{N}}$ *strongly converges in* V *to* u.

The following result is about the speed of convergence:

Theorem 2.4. *If the spaces* $V_1, V_2, \ldots, V_d$ *are finite dimensional, the convergence rate of the algorithm is exponential, in other words, there exists* $C, \sigma > 0$ *such that for all* $n \in \mathbb{N}^*$

$$\|u - u_n\|_V \leq Ce^{-\sigma n}$$

The constant C can be estimated by $\|u\|_V$ and the constant σ depends on the dimensions of the spaces $V_1, V_2, \ldots, V_d$.

Another important result obtained by Cancès, Ehrlacher and Lelièvre [CEL12] is that if we suppose in addition that

(A6) There exists $\beta, \gamma > 0$ such that for all $(r^1, r^2) \in V_1 \times V_2$, with V_1 and V_2 two Hilbert spaces.

$$\beta \|r^1\|_{V_1} \|r^2\|_{V_2} \leq \|r^1 \otimes r^2\| \leq \gamma \|r^1\|_{V_1} \|r^2\|_{V_2}$$

then it is not necessary to obtain the global minimum of (2.8) to ensure the convergence of the greedy algorithm. We note that this result is proved when it is considered a product of only two Hilbert spaces.

Theorem 2.5. *Let us suppose that we are in the case of only two Hilbert spaces* V_1 *and* V_2 *and that the assumptions (A1)-(A6) are satisfied. Then, if at each iteration* $n \in \mathbb{N}$, *the pair* $(r^1, r^2) \in V_1 \times V_2$ *is chosen to be a local minimum of* (2.8), *such that* $\mathcal{E}(u_n) < \mathcal{E}(u_{n-1})$, *then* $(u_n)_{n\in\mathbb{N}^*}$ *still converges strongly in* V *towards* u *the solution of* (2.6). *Besides, if the Hilbert spaces* V_1 *and* V_2 *are finite dimensional, the rate of convergence of the algorithm is still exponential in* n, *i.e., there exists* $C, \sigma > 0$ *such that for all* $n \in \mathbb{N}^*$

$$\|u - u_n\|_V \leq Ce^{-\sigma n}$$

2.5 The Proper Generalized Decomposition for the approximation of a square-integrable function

In order to show the implementation that we use for the algorithm (2.7), let us present the simple problem of approximating a square-integrable function f by a sum of tensor products. Mathematically, we consider the spaces $V = L^2(\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d)$, $V_{x_i} = L^2(\Omega_i)$ for $i = 1, \ldots, d$, where $\Omega_i \subset \mathbb{R}$ is a bounded domain for i such that $1 \leq i \leq d$. We recall that we are looking for a separated representation $f = \sum_{k \geq 1} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d$. So, let us consider the following minimization problem:

$$\text{Find } u \in L^2(\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d) \text{ such that } u = \arg\min_{v \in L^2} \left(\frac{1}{2} \int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} v^2 - \int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} v f \right) \tag{2.16}$$

whose solution is obviously $u = f$. In this context, the algorithm of the PGD (2.7) can be rewritten as follows:

Iterate for all $n \geq 1$: Find $(r_n^1, r_n^2, \ldots, r_n^d) \in V_{x_1} \times V_{x_2} \times \ldots \times V_{x_d}$ such that $(r_n^1, r_n^2, \ldots, r_n^d)$ belongs to

$$\begin{aligned} \underset{r^1 \in L^2(\Omega_1), \ldots, r^d \in L^2(\Omega_d)}{\operatorname{argmin}} \quad & \frac{1}{2} \int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} \left| \sum_{k=1}^{n-1} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d + r^1 \otimes r^2 \otimes \ldots \otimes r^d \right|^2 \\ & - \int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} \left(\sum_{k=1}^{n-1} r_k^1 \otimes r_k^2 \otimes \ldots r_k^d + r^1 \otimes r^2 \otimes \ldots \otimes r^d \right) f, \end{aligned} \tag{2.17}$$

As proposed in [BLM09], instead of solving the problem (2.17), we will determine the solutions of the Euler equation for (2.17). Notice that, in general, the solutions of the Euler equation are not necessarily the solutions of the minimization problem, given the nonlinearity of the tensor product space $L^2(\Omega_1) \otimes L^2(\Omega_2) \otimes \ldots \otimes L^2(\Omega_d)$.

The Euler equation for (2.17) has the following form:

Find $(r_n^1, r_n^2, \ldots, r_n^d) \in L^2(\Omega_1) \times L^2(\Omega_2) \times \ldots \times L^2(\Omega_d)$ such that for any functions $(r^1, r^2, \ldots, r^d) \in L^2(\Omega_1) \times L^2(\Omega_2) \times \ldots \times L^2(\Omega_d)$, we have

$$\begin{aligned} & \int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} (r_n^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d) \left(r^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d + r_n^1 \otimes r^2 \otimes \ldots \otimes r_n^d + \ldots + r_n^1 \otimes r_n^2 \otimes \ldots \otimes r^d \right) \\ & = \int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} f_{n-1} \left(r^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d + r_n^1 \otimes r^2 \otimes \ldots \otimes r_n^d + \ldots + r_n^1 \otimes r_n^2 \otimes \ldots \otimes r^d \right) \end{aligned} \tag{2.18}$$

where $f_{n-1} = f - \sum_{k=1}^{n-1} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d$.

The equation (2.18) can be written equivalently as

$$\begin{cases} \|r_n^2\|^2 \|r_n^3\|^2 \ldots \|r_n^d\|^2 \; r_n^1 = \int_{\Omega_2 \times \Omega_3 \times \ldots \times \Omega_d} \left(r_n^2 \otimes \ldots \otimes r_n^d \right) \; f_{n-1}, \\ \|r_n^1\|^2 \|r_n^3\|^2 \|r_n^4\|^2 \ldots \|r_n^d\|^2 \; r_n^2 = \int_{\Omega_1 \times \Omega_3 \times \Omega_4 \ldots \times \Omega_d} \left(r_n^1 \otimes r_n^3 \otimes r_n^4 \otimes \ldots \otimes r_n^d \right) \; f_{n-1}, \\ \quad \vdots \\ \|r_n^1\|^2 \|r_n^2\|^2 \|r_n^3\|^2 \ldots \|r_n^{d-1}\|^2 \; r_n^d = \int_{\Omega_1 \times \Omega_2 \times \Omega_3 \ldots \times \Omega_{d-1}} \left(r_n^1 \otimes r_n^2 \otimes r_n^3 \otimes \ldots \otimes r_n^{d-1} \right) \; f_{n-1} \quad , \end{cases} \tag{2.19}$$

where $\|r_n^i\|^2$ denotes the L^2 norm: $\|r_n^i\|^2 = \int_{\Omega_i} |r_n^i|^2$.

The system (2.19) is a non linear coupled system of equations which can be solved by a fixed point procedure as proposed in [ACKM06]. Choose $(r_n^{1,(0)}, r_n^{2,(0)}, \ldots, r^{d,(0)}) \in L^2(\Omega_1) \times L^2(\Omega_2) \times \ldots \times L^2(\Omega_d)$, and at iteration $k \geq 0$, compute $(r_n^{1,(k)}, r_n^{2,(k)}, \ldots, r^{d,(k)}) \in L^2(\Omega_1) \times L^2(\Omega_2) \times \ldots \times L^2(\Omega_d)$ which is the solution to

$$\begin{cases} \|r_n^{2,(k)}\|^2 \|r_n^{3,(k)}\|^2 \ldots \|r_n^{d,(k)}\|^2 \; r_n^{1,(k+1)} = \int_{\Omega_2 \times \Omega_3 \times \ldots \times \Omega_d} \left(r_n^{2,(k)} \otimes \ldots \otimes r_n^{d,(k)} \right) \; f_{n-1}, \\ \|r_n^{1,(k+1)}\|^2 \|r_n^{3,(k)}\|^2 \|r_n^{4,(k)}\|^2 \ldots \|r_n^{d,(k)}\|^2 \; r_n^{2,(k+1)} = \int_{\Omega_1 \times \Omega_3 \times \Omega_4 \ldots \times \Omega_d} \left(r_n^{1,(k+1)} \otimes r_n^{3,(k)} \otimes r_n^{4,(k)} \otimes \ldots \otimes r_n^{d,(k)} \right) \; f_{n-1}, \\ \quad \vdots \\ \|r_n^{1,(k+1)}\|^2 \|r_n^{2,(k+1)}|^2 \|r_n^{3,(k+1)}\|^2 \ldots \|r_n^{d-1,(k+1)}\|^2 \; r_n^{d,(k+1)} \\ = \int_{\Omega_1 \times \Omega_2 \times \Omega_3 \ldots \times \Omega_{d-1}} \left(r_n^{1,(k+1)} \otimes r_n^{2,(k+1)} \otimes r_n^{3,(k+1)} \otimes \ldots \otimes r_n^{d-1,(k+1)} \right) \; f_{n-1}, \end{cases} \tag{2.20}$$

until convergence is reached.

It is important to note that we start with a linear problem (2.16) with exponential complexity with respect to the dimension, and we obtain at the end a nonlinear problem (2.19) with a linear complexity with respect to the dimension at each iteration . This is a general feature of the PGD method: the curse of dimensionality is circumvented, but the linearity of the original problem is lost because the space of tensor products is non-linear.

In the two-dimensional case ($d = 2$), the algorithm given by (2.17) is related to the Singular Value Decomposition (or rank one decomposition), as it is explained in [BLM09]. In this case, the solutions of the variational problem (2.17) are exactly the solutions to the Euler equation (2.18) that verify the second-order optimality conditions (See Section 2.3). This property does not hold in a d-dimensional framework with $d \geq 3$.

2.6 The Proper Generalized Decomposition in the case of the Poisson problem

In this section we present the application of the Proper Generalized Decomposition for solving high-dimensional PDEs through the use of the greedy algorithms.

As in [BLM09], we take the example of the Poisson problem. Let us consider a function $f \in L^2(\Omega)$, where $\Omega = \Omega_1 \times \ldots \times \Omega_d$ with $\Omega_i \subset \mathbb{R}$ are bounded domains for all $i = 1, \ldots d$. Thus, we define the following homogeneous multivariate Dirichlet Poisson problem:

$$\text{Find } u \in H_1^0(\Omega) \text{ such that } \begin{cases} -\Delta u = f \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases} \tag{2.21}$$

Using Lax Milgram's theorem, (2.21) is equivalent to the following minimization problem

$$\text{Find } u \in H_1^0(\Omega) \text{ such that } u = \underset{v \in H_1^0(\Omega)}{\operatorname{argmin}} \left(\frac{1}{2} \int_\Omega |\nabla v|^2 - \int_\Omega f v \right). \tag{2.22}$$

Moreover, we note that the problem (2.22) is equivalent to

$$\text{Find } u \in H_1^0(\Omega) \text{ such that } u = \underset{v \in H_1^0(\Omega)}{\operatorname{argmin}} \int_\Omega |\nabla (v - u)|^2, \tag{2.23}$$

where u is the solution of the problem (2.22).

In particular, in the case of the Poisson problem (2.21), the greedy algorithm writes:

Find $\left(r_n^1 \otimes r_n^2 \otimes \times \otimes r_n^d\right) \in H_1^0(\Omega_1) \times H_1^0(\Omega_2) \times \ldots \times H_d^0(\Omega)$ such that

$$\left(r_n^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d\right) \in \underset{r^1 \in H_1^0(\Omega_1), r^2 \in H_1^0(\Omega_2), r^d \in H_1^0(\Omega_d)}{\operatorname{argmin}} \frac{1}{2} \int_\Omega |\nabla \left(r^1 \otimes r^2 \otimes \ldots \otimes r^d\right)|^2 - \int_\Omega f_{n-1} r_n^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d \tag{2.24}$$

where $f_{n-1} = f + \Delta \left(\sum_{k=1}^{n-1} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d \right)$.

The Euler-Lagrange equation related to the minimization problem (2.24) is: Find $(r_n^1, r_n^2, \ldots, r_n^d) \in H_1^0(\Omega_1) \times H_1^0(\Omega_2) \times \ldots \times H_1^0(\Omega_d)$ such that for any function $(r^1, r^2, \ldots, r^d) \in H_1^0(\Omega_1) \times H_1^0(\Omega_2) \times H_1^0(\Omega_d)$

$$\begin{aligned} \int_\Omega \nabla \left(r_n^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d\right) \cdot \nabla \left(r^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d + r_n^1 \otimes r^2 \otimes \ldots \otimes r_n^d + \ldots + r_n^1 \otimes r_n^2 \otimes \ldots \otimes r^d\right) = \\ = \int_\Omega f_{n-1} \left(r^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d + r_n^1 \otimes r^2 \otimes \ldots \otimes r_n^d + \ldots + r_n^1 \otimes r_n^2 \otimes \ldots \otimes r^d\right) \end{aligned} \tag{2.25}$$

This Euler-Lagrange equation can be written as follows

$$
\begin{cases}
-\left(\|r_n^2\|_{L^2}^2\dots\|r_n^d\|_{L^2}^2\right)(r_n^1)'' + \sum_{i=2}^d \prod_{j=2,j\neq i}^d \|r_n^j\|_{L^2}^2 r_n^1 = \int_{\Omega_2\times\ldots\times\Omega_d} f_{n-1} r_n^2 \otimes \ldots \otimes r_n^d \\
-\left(\|r_n^1\|_{L^2}^2\|r_n^3\|_{L^2}^2\dots\|r_n^d\|_{L^2}^2\right)(r_n^2)'' + \sum_{i=1,i\neq 2}^d \prod_{j=1,j\neq 2,j\neq i}^d \|r_n^j\|_{L^2}^2 r_n^2 = \int_{\Omega_1\times\Omega_3\ldots\times\Omega_d} f_{n-1} r_n^1 \otimes r_n^3 \ldots \otimes r_n^d \\
\vdots \\
-\left(\|r_n^1\|_{L^2}^2\dots\|r_n^{d-1}\|_{L^2}^2\right)\left(r_n^d\right)'' + \sum_{i=1,i\neq d}^d \prod_{j=1,j\neq d,j\neq i}^d \|r_n^j\|_{L^2}^2 r_n^d = \int_{\Omega_1\times\ldots\times\Omega_{d-1}} f_{n-1} r_n^1 \otimes \ldots \otimes r_n^{d-1}
\end{cases}
\tag{2.26}
$$

We observe that the Euler equation (2.25) is not equivalent to the original problem (2.24) because of the nonlinearity of the tensor product space $H_0^1(\Omega_1)\otimes\ldots\otimes H_0^1(\Omega_d)$

We remark that calculating high-dimensional integrals as:

$$
\int_{\Omega_1\times\Omega_2\ldots\times\Omega_d} f(x_1,x_2,\ldots,x_d)\, r_n^1(x_1)\otimes\ldots\otimes r_n^d \;\; dx_1 dx_2 \ldots dx_d \tag{2.27}
$$

can be very costly. The idea that we use to overcome this practical obstacle is to approximate, in a preliminary step, the data (the function f) by a sum of tensor products using, for example, the approach presented in Section 2.5 and then to use Fubini's rule to compute the term (2.27).

As in the case of the approximation of a square-integrable function, we start from a linear problem with exponential complexity with respect to the dimension d as (2.21) and we end up with a nonlinear problem with linear complexity with respect to the dimension d given by (2.26).

We observe that (2.26) is a nonlinear coupled system of low-dimensional Poisson equations, which may be solved by a simple fixed point procedure as follows: At iteration $l\leq 1$, find the functions $(r_n^{(1),l}, r_n^{(2),l},\ldots,r_n^{(d),l})\in H_1^0(\Omega_1)\times H_1^0(\Omega_2)\ldots\times H_1^0(\Omega_d)$ solution to

$$
\begin{cases}
-\left(\|r_n^{2,l-1}\|_{L^2}^2\dots\|r_n^{d,l-1}\|_{L^2}^2\right)\left(r_n^{1,l}\right)'' + \sum_{i=2}^d \prod_{j=2,j\neq i}^d \|r_n^{j,l-1}\|_{L^2}^2 r_n^{1,l} = \int_{\Omega_2\times\ldots\times\Omega_d} f_{n-1} r_n^{2,l-1} \otimes \ldots \otimes r_n^{d,l-1} \\
-\left(\|r_n^{1,l}\|_{L^2}^2\|r_n^{3,l-1}\|_{L^2}^2\dots\|r_n^{d,l-1}\|_{L^2}^2\right)\left(r_n^{2,l}\right)'' + \sum_{i=1,i\neq 2}^d \prod_{j=1,j\neq 2,j\neq i}^d \|r_n^{j,l-1}\|_{L^2}^2 r_n^{2,l} \\
= \int_{\Omega_1\times\Omega_3\ldots\times\Omega_d} f_{n-1} r_n^{1,l} \otimes r_n^{3,l-1} \ldots \otimes r_n^{d,l-1} \\
\vdots \\
-\left(\|r_n^{1,l}\|_{L^2}^2\dots\|r_n^{d-1,l}\|_{L^2}^2\right)\left(r_n^{d,l-1}\right)'' + \sum_{i=1,i\neq d}^d \prod_{j=1,j\neq d,j\neq i}^d \|r_n^{j,l}\|_{L^2}^2 r_n^{d,l-1} = \int_{\Omega_1\times\ldots\times\Omega_{d-1}} f_{n-1} r_n^{1,l} \otimes \ldots \otimes r_n^{d-1,l}
\end{cases}
\tag{2.28}
$$

We remark that the number of degrees of freedom needed to obtain the solution u_n in the form (2.5) is $n\sum_{i=1}^d N_i$ where N_i is the number of discretization points in the dimension i. So, the number of unknowns at each iteration of the algorithm (2.7) grows linearly with respect to the dimension and this allows an important reduction in the number of degrees of freedom with respect to the classical approach. At each iteration of the greedy algorithm, the computational cost is given by the solution of a low-dimensional nonlinear problem which is deduced from the high-dimensional linear problem as in the example given in the previous section 2.6.

Convergence results for the Poisson problem

Le Bris, Lelièvre and Maday in [BLM09] prove that the PGD (2.24) converges when the PGA and OGA are used to solve the Poisson problem.

Theorem 2.6. *Let us assume that the assumptions (A1), (A2) and (A3) are verified and that the functional $\mathcal{E}$ has the form* (2.12)*. Then, for a function $u \in \mathcal{A}^1$, there exists a constant $C > 0$ such that*

$$\|u_n - u\|_V \leq C n^{-1/6}, \tag{2.29}$$

for all $n \in \mathbb{N}^*$.

As mentioned above, the convergence rate factor of $\frac{1}{6}$ can be improved to $\frac{11}{62}$ and that the constant C depends on the norm $\|u\|_{\mathcal{A}^1}$.

An extension of this result is proposed by Figueroa and Sully in [FS12] in the case of an arbitrary diffusion coefficient satisfying a uniform ellipticity assumption.

Remark 2.2. *A full characterization of the set $\mathcal{A}_1(\mathcal{D})$ is not clear. Le Bris, Lelièvre and Maday in [BLM09] give the following characterization of this set in terms of standard Sobolev spaces for the Poisson problem* (2.21)

$$\textit{For any } m > 1 + \frac{d}{2}, \quad H^m(\Omega) \cap H^0_1(\Omega) \subset \mathcal{A}_1(\mathcal{D})$$

More characterizations for the set $\mathcal{A}_1(\mathcal{D})$ are obtained in [FS12] for more general operators.

3

Approximation of a Put payoff function using the Proper Generalized Decomposition

In this section, we discuss the implementation of the Proper Generalized Decomposition defined by (2.7) in the case of the approximation of a put payoff function (that is, a square-integrable function) by a sum of tensor products as we presented in Section 2.5. We will then provide numerical examples of the application of this approach. This particular case has the advantage of being an easy example to understand the implementation of the algorithm (2.7). Moreover, this procedure is useful in a preliminary step to approximate the initial condition of the Black-Scholes PDE (see Section 4.1.5), namely in order to get a separated representation of the payoff function.

3.1 Separated representation of a Put payoff

In this section we will apply the algorithm (2.17) to obtain an approximation of the payoff of a basket put option. This implies that we take $f(x_1, \dots, x_d) = \left(K - \frac{1}{d}\sum_{i=1}^{d} x_i\right)_+$ in the algorithm (2.17). For the practical implementation of the greedy algorithm, we need to introduce the space discretization. In practice, the spaces $V_{x_i}^{\Delta x}$ for $i = 1, \dots, d$ that are used to discretize $L^2(\Omega_i)$ with $\Omega_i = (0, 1)$ for $i = 1, \dots d$ are the P1 finite elements on a uniform mesh with space step Δx. The number $N = \frac{1}{\Delta x}$ is the number of intervals in each direction. For each k, we discretize the functions r_k^i for $i = 1, \dots, d$ that appear in the approximation of the solution given by the expression (2.18) as follows:

$$r_k^{i,\Delta x}(x_i) = \sum_{j=0}^{N} r_k^{i,j} \phi_j(x_i), \quad r_k^{i,j} \in \mathbb{R}, \ \forall j, k, \tag{3.1}$$

where $\phi_i(x) = \phi\left(\frac{x - x_i}{\Delta x}\right)$ with $\phi(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| > 0. \end{cases}$

This type of discretization and its generalization to the d-dimensional case will be used for all the numerical simulations of this work.

Figure 3.1 shows how the algorithm approximates the basket put payoff in a two-dimensional framework ($d = 2$). We observe that, as the number of iterations of the greedy algorithm increases, the approximation of the function $f(x_1, x_2)$ improves.

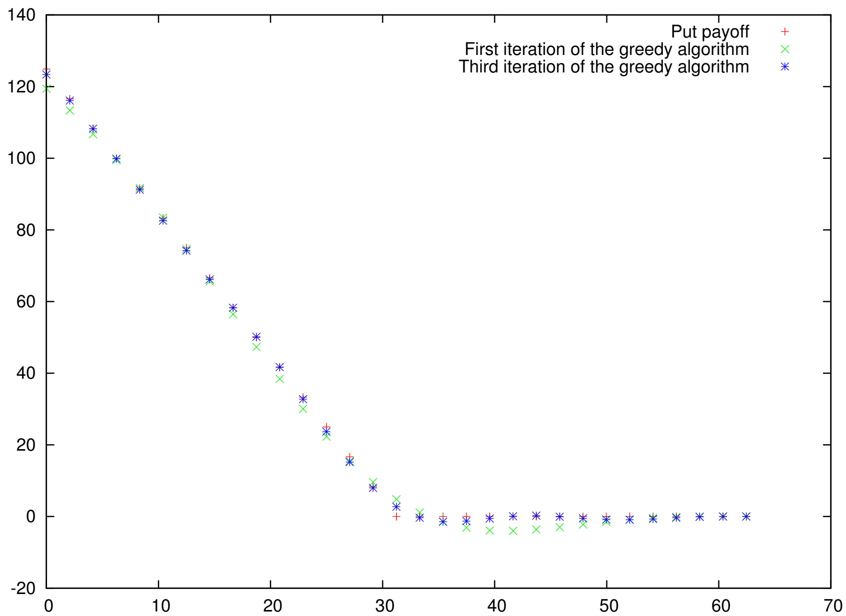


Fig. 3.1. Basket put option with two assets. We consider here the intersection between the surface of prices and the plane $S_1 = S_2$. To obtain this approximation we take 31 points of discretization per dimension (N= 30). In this figure, we show the approximation given after the first and third iteration of the algorithm.

In order to study numerically the problem of approximating a square-integrable function, we discuss about the fixed point procedure, the criteria of convergence of the algorithm (2.7) and the numerical integration techniques used in our numerical experiments.

3.2 Fixed point procedure

In what follows, we are going to study the convergence of the fixed point method (2.20) from a numerical point of view. This method allows us to find the solution of the Euler equation represented by the non-linear system of equations (2.19). In order to analyze the fixed point procedure used in practice in this work, we will propose three different stopping criteria for two different norms: L^∞ and L^2.

The first criterion consists in simply fixing arbitrarily the number of iterations. This arbitrary number of iterations can be determined, for instance, by iterating the algorithm many times and thus deduce the needed number by experience. The advantage of such a method is that we do not have to compute any norm of the error at each step of the fixed point procedure (2.20) and so we can reduce the computational time. Nevertheless, the inconvenient is that we cannot predict if the chosen number is optimal.

A second criterion, called the residues method, consists in determining the convergence of the fixed point method by calculating the following norm (L^2 or L^∞):

$$\frac{\|r_n^{1,k+1} \otimes r_n^{2,k+1} \otimes ... \otimes r_n^{d,k+1} - r_n^{1,k} \otimes r_n^{2,k} \otimes ... \otimes r_n^{d,k}\|}{\|r_n^{1,k} \otimes r_n^{2,k} \otimes ... \otimes r_n^{d,k}\|} \tag{3.2}$$

where n is fixed and represents the PGD iteration and k is the fixed point iteration.

The third implemented method compares the solutions obtained in two consecutive iterations of the fixed point method fixing one direction, by computing the following norm in L^2 or a L^∞:

$$\frac{\|r_n^{i,k+1} - r_n^{i,k}\|}{\|r_n^{i,k}\|} \tag{3.3}$$

Let us now, to compare the last two criteria according to the number of iterations of the fixed point procedure (2.20). Figure 3.2 illustrates the evolution of the number of iterations needed for the convergence of the fixed point procedure when we use the expressions given by (3.2) and (3.3) for a basket put option function with 5 assets. For practical considerations, we consider a maximum number of 30 fixed point iterations by dimension, so for each PGD iteration we compute a maximum of 150 fixed point iterations in a five-dimensional framework. We observe that for this example, the criterion (3.3) yields to less iterations for the fixed point procedure compared with the criterion (3.2).

Nevertheless, we cannot say that we should always use the criterion (3.3), as in the case of a basket put option with 6 assets, the criterion (3.3) needs more iterations to converge compared with the criterion (3.2) as illustrated in Figure 3.3.

As a conclusion, we can say that for our numerical experiences, we have chosen an arbitrary number of iterations for the fixed point method because neither of the two criteria (3.3) and (3.2) does not seem to influence the convergence of the PGD algorithm.

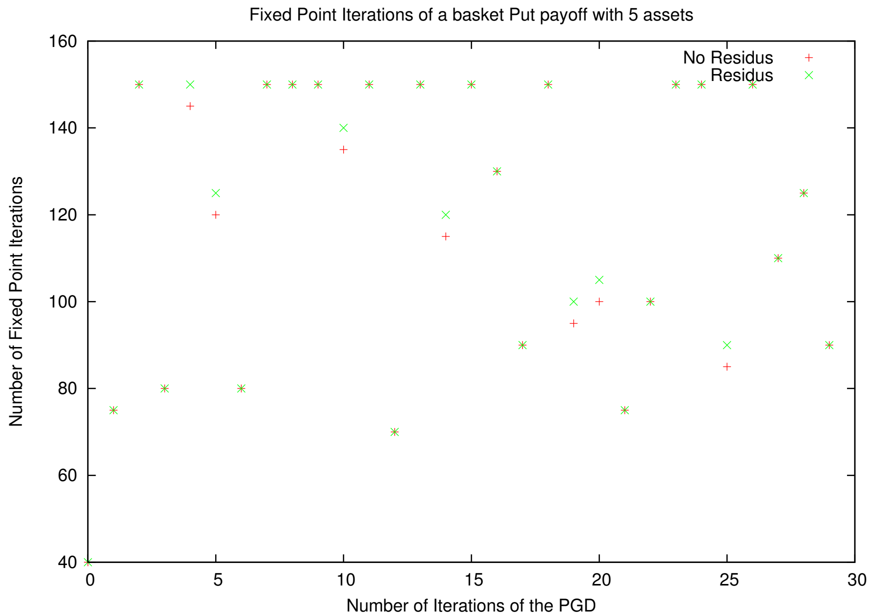


Fig. 3.2. Comparison between the number of fixed point iterations obtained at each PGD iteration according to the criteria (3.2) and (3.3).

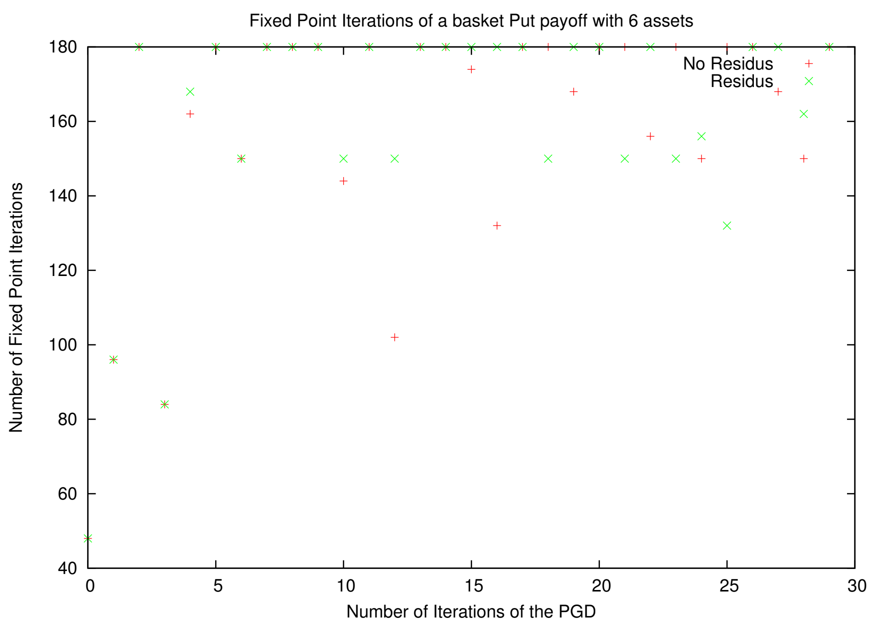


Fig. 3.3. Comparison between the number of fixed point iterations obtained at each PGD iteration according to the criteria (3.2) and (3.3).

3.3 Initial condition for the fixed point procedure and for the convergence of the Proper Generalized Decomposition method

A choice that can influence the convergence of the fixed point procedure is the initial condition. To analyze this point, we study numerically two cases: a constant initial condition and a random initial condition.

The constant initial condition is equal to the matrix $\frac{\|u(x_1,\ldots,x_d)-\sum_{k=1}^{n-1} r_k^1\otimes\ldots\otimes r_k^d\|}{\|\mathbf{1}\|}\mathbf{1}$, where $\mathbf{1}$ is the matrix such that $\mathbf{1}_{i,j} = 1$ for all i, j, whereas the random initial condition is given by vectors whose coordinates are uniform random variables in the set $[-1, 1]$. We renormalize this vector to obtain a norm equal to the current error $\|u(x_1, \ldots, x_d) - \sum_{k=1}^{n-1} r_k^1 \otimes \ldots \otimes r_k^d\|$, where n is the current PGD iteration and u is the basket put payoff with d assets.

Figures 3.4 and 3.5 compare the number of fixed point iterations needed under the criterion (3.3) depending on the choice of a random or constant initial condition. Using these figures, we outline that it is not clear if it would be better to begin the fixed point procedure with a random or a constant initial condition.

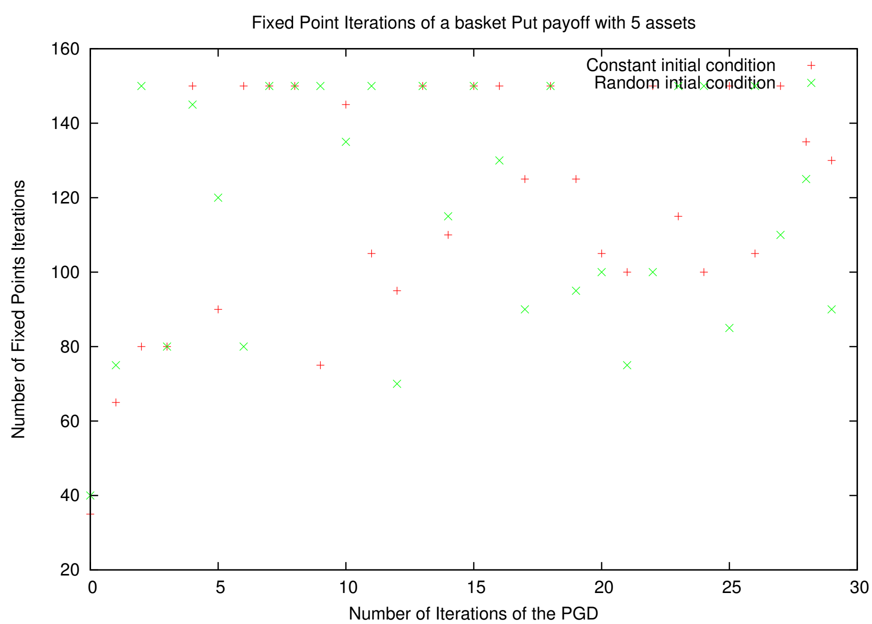


Fig. 3.4. Comparison of the number of fixed point iterations needed to obtain convergence according to the election of a random or a constant initial condition in a 5 dimensional framework.

The choice of a random or a constant initial condition for the fixed point procedure has to be analyzed as well from the convergence of the algorithm (2.7) to the basket put option function as the number of the PGD iterations increases.

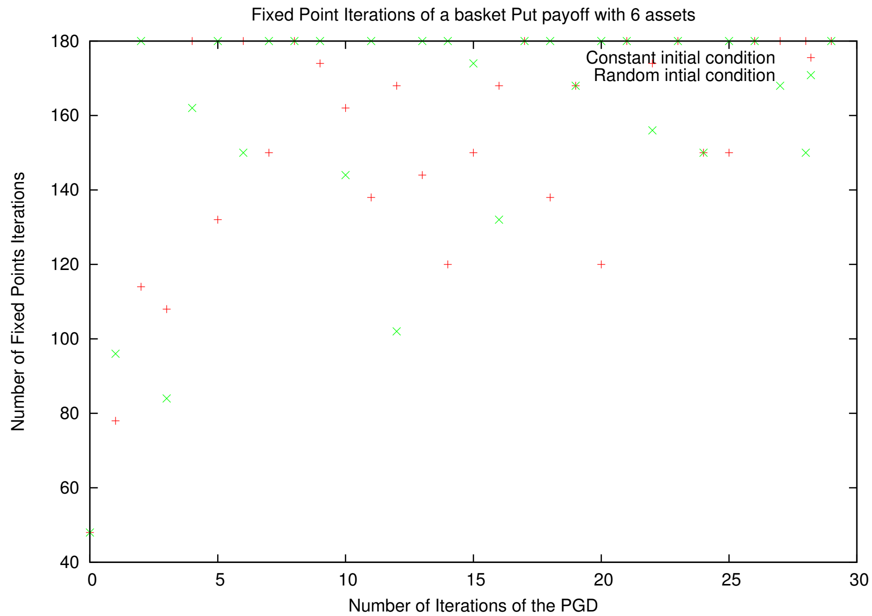


Fig. 3.5. Comparison of the number of fixed point iterations needed to obtain convergence according to the election of a random or a constant initial condition in a 6 dimensional framework.

Figure 3.6 and figure 3.7 show a comparison between the convergence curves for a basket put option approximation in the case of 5 and 6 assets respectively obtained with a random and a constant initial condition. We note that in this case, using a random initial condition leads to a faster convergence. The experience obtained by this type of observations led us to consider random initial conditions for the fixed point procedure in our numerical examples.

3.4 Criteria of convergence used in practice

In this part, we discuss about the criteria that we used in practice to establish the convergence of the PGD algorithm (2.17). We note that the computation of

$$\|f(x_1,\ldots,x_d) - \sum_{k=1}^{n} r_k^1 \otimes \ldots \otimes r_k^d\|, \tag{3.4}$$

where f is the square-integrable function to approximate, involves a high-dimensional integral. We give more details on the treatment of terms of type (3.4) in Section 3.5.

A possible criterion would be to verify if $\|r_n^1 \otimes \ldots \otimes r_n^d\|$ is small compared with $\|\sum_{k=1}^{n-1} r_k^1 \otimes \ldots \otimes r_k^d\|$. In [BLM09], the authors note that using this criterion can lead to missing a term with a large contribution in terms of norm. They show that in the fixed point procedure (2.20), an optimal term

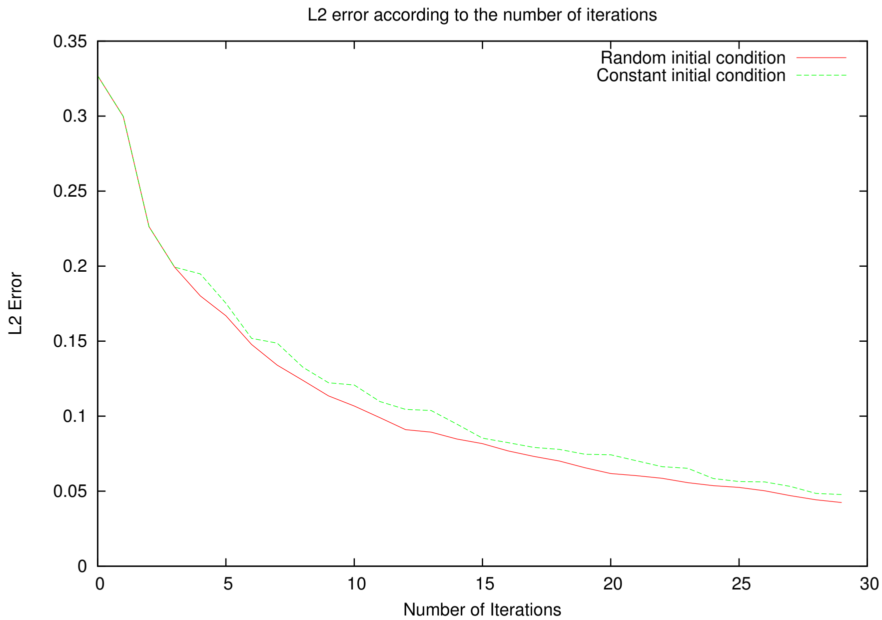


Fig. 3.6. Comparison of the convergence curves for a basket put option with 5 assets according to the use of a random or a constant initial condition

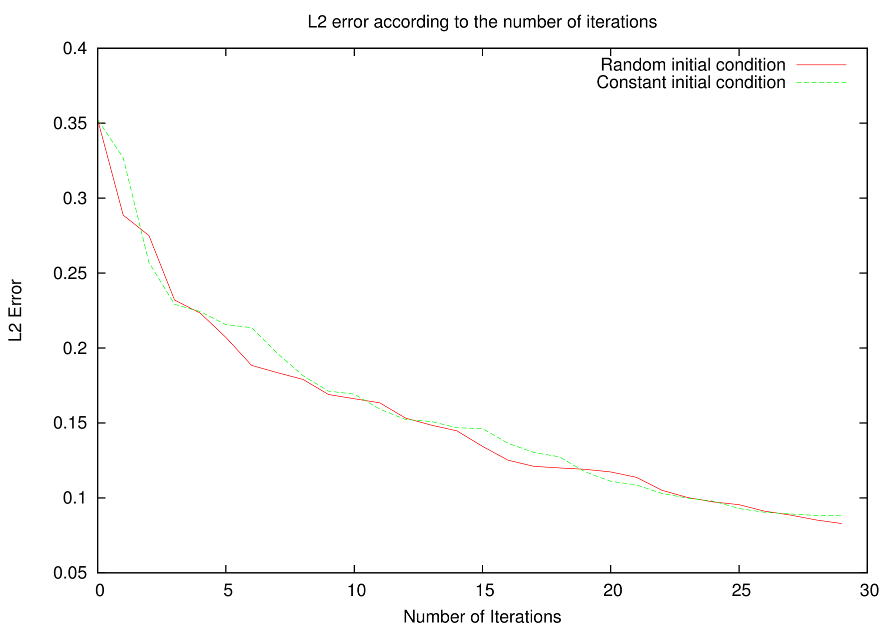


Fig. 3.7. Comparison of the convergence curves for a basket put option with 6 assets according to the use of a random or a constant initial condition

can be missed if the initial condition has a zero component on the eigenspace associated to this term. They also note that it is possible that the optimal terms $r_n^1, \ldots, r_n^d$ obtained when solving the Euler equation are picked in an order not appropriate for computational efficiency. Moreover, they remark that the relaxation step performed in the orthogonal version of the PGD does not solve any of these difficulties.

As the computation of the term (3.4) in dimension d, when d is large, is not possible for storage reasons, from a practical point of view, we use a number of iterations for the algorithm (2.17) defined a priori. This number of iterations can be estimated using many tries with the algorithm measuring at the end the norm of the error (3.4). We remark that even when the dimension d is not large, the computation of the error term (3.4) can take a lot of time.

In our numerical experiments, we calculate the support of the payoff function and we store the points obtained in order to use them in the next iterations of the algorithm. In this manner, a backtracking algorithm is used only one time in the computation of the approximation of the function f.

3.5 Numerical integration

In order to reduce the computational time needed to calculate the integrals presented in the system of equations (2.19), we use the specific form of the payoff function in order to compute, in a preliminary step, the points that belong to the support of this function. Thus, when we calculate numerically the integral term

$$\int_{\Omega_1 \times \Omega_2 \times \ldots \times \Omega_d} \left(r_n^1 \otimes r_n^2 \otimes \ldots \otimes r_n^d \right) \, f \, dx_1 dx_2 \ldots dx_d \tag{3.5}$$

in (2.19), we do not need to pass through the points where the function vanishes. In practice, in order to describe the support of this payoff function, we use a backtracking algorithm. This type of algorithm consists in constructing candidates sequentially and neglecting them when they do not verify the conditions required as a solution, in this case to belong to the support of the payoff.

Specifically, we consider the function $(K - \frac{1}{d}\sum_{i=1}^{d} S_i)_+$ with $S_i \in [0, S^{max}]$ for all $i = 1, \ldots d$. Introducing the parameter N to discretize the interval $[0, S^{max}]$, we obtain that $S_j^i = x_j^i S^{max}$, with $x_j^i = \frac{(j-1)}{N}$, $j = 1, \ldots, N+1$. This leads us to look for those x^i, $i = 1, \ldots d$ such as

$$dK - S^{max} \sum_{i=1}^{d} x_j^i > 0$$

for $j = 1, \ldots, N+1$. In order to do that, we need to find the combinations of $(i_1, ..., i_d) \in \mathbb{N}^d$, such that:

$$\frac{i_1}{N} + ... + \frac{i_d}{N} \leq d\frac{K}{S^{max}}$$

From a point $(i_1, i_2 \ldots, i_d) \in \mathbb{N}^d$ that belongs to the support of the function f, we increase i_1 by 1 and we check if this new point $(i_1 + 1, i_2 \ldots, i_d)$ is in the support of f. If this is the case, we continue incrementing i_1 until the considered point is no longer in the support of f and then we pass to the following coordinate.

For instance, in a 5-dimensional case, the computational time is reduced by a factor of $\frac{4}{5}$ by taking into account the support of the payoff in the computations of the integral term (3.5).

Remark 3.1. *Concerning the computational time, we note that if the dimension increases, one iteration of the algorithm takes more time to be computed because the number of equations in the Euler system* (2.19) *increases linearly with respect to the dimension. The integral terms of type* (3.5) *also demand more time of execution because the domain has a new variable.*

3.6 Numerical results

Figure 3.8 shows that, as the dimension increases, the number of iterations needed to obtain the convergence increases as well.

We also provide in Table 3.1 the number of iterations needed in order to obtain a relative error of 10^{-5} when we consider 11 points of discretization per dimension ($N = 10$). The relative error calculated is the discrete L^2 error

$$e_n = \frac{\sqrt{\frac{1}{N}\sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\ldots\sum_{i_d=1}^{N}\left(f(x_{i_1}, x_{i_2}, \ldots, x_{i_d}) - u^n(x_{i_1}, x_{i_2}, \ldots, x_{i_d})\right)^2}}{\sqrt{\frac{1}{N}\sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\ldots\sum_{i_d=1}^{N} f(x_1, \ldots, x_d)^2}} \tag{3.6}$$

where $u^n(x_1, x_2, \ldots, x_d) = \sum_{k=1}^{n} r_k^1 \otimes r_k^2 \otimes \ldots \otimes r_k^d$ is the solution obtained with the PGD algorithm at the iteration n. It is to be outlined that computing (3.6) is more costly than the PGD method itself.

Notice that the full tensor product approximation would require $11^8 \simeq 2.10^8$ degrees of freedom in an 8-dimensional case, whereas we obtain $3974 \times 8 \times 11 \simeq 350000$ degrees of freedom. For the evaluation of this number, we used the fact that at each iteration of the algorithm we get 8 functions that are determined by 11 discretization points.

We can also study the convergence of the algorithm with respect to the L^∞ norm. Figure 3.9 illustrates the convergence curves for the L^2 and L^∞ norm when we approximate a basket put option with 3 assets.

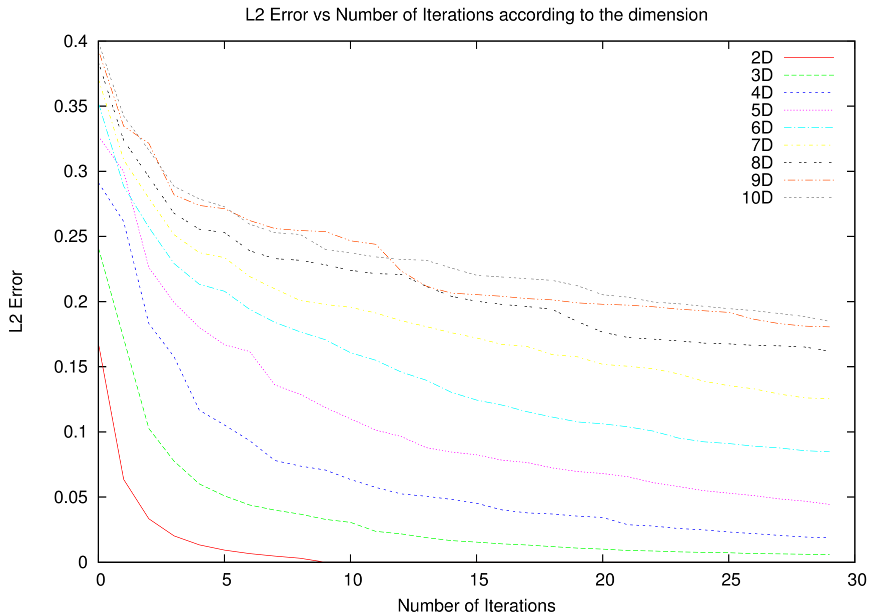


Fig. 3.8. Convergence curves for the approximation of a basket put payoff by a sum of tensor products. We observe that if the dimension increases, the number of iterations needed to obtain the convergence increases as well. The error calculated is given by (3.6).

Dimension	Number of iterations
1	1
2	2
3	10
4	22
5	101
6	228
7	1077
8	3974

Table 3.1. Number of iterations needed to obtain a relative error of 10^{-5} when we take 11 discretization points per dimension

3.7 Mass lumping technique

Until here, all the numerical results presented were obtained by computing the term (3.5) using the following approximation:

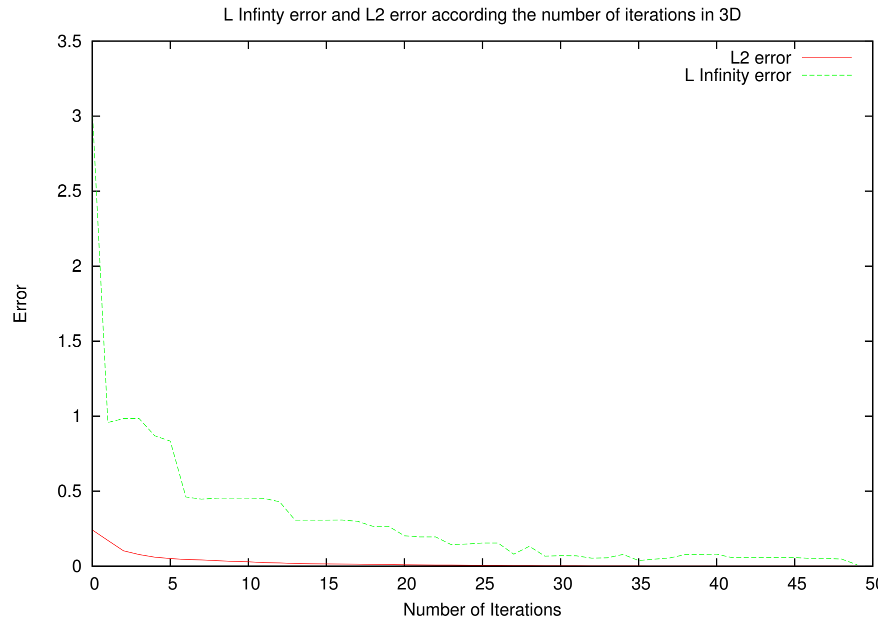


Fig. 3.9. Convergence curves in the L^2 and L^∞ norm for the approximation of a basket put payoff with 3 assets by a sum of tensor products.

$$\begin{aligned}\int_{\Omega_1\times\ldots\times\Omega_d}\left(r_n^1\otimes\ldots\otimes r_n^d\right)f(x_1,\ldots,x_d)dx_1\ldots dx_d &\approx \int_{\Omega_1\times\ldots\times\Omega_d}\sum_{i_1=0}^{N}\ldots\sum_{i_d=0}^{N} f(x_1^{i_1},\ldots,x_d^{i_d})\phi_{i_1}(x_1)\ldots\phi_{i_d}(x_d)\\ &\quad\left(\sum_{j_1=0}^{N} r_1^{j_1}\phi_{j_1}(x_1)\right)\ldots\left(\sum_{j_d=0}^{N} r_1^{j_d}\phi_{j_d}(x_d)\right)dx_1\ldots dx_d\\ &= \sum_{i_1=0}^{N}\ldots\sum_{i_d=0}^{N} f(x_1^{i_1},\ldots,x_d^{i_d})\\ &\quad\left(\sum_{j_1=i_1-1}^{i_1+1} r_n^{1,j_1}\int_{\Omega_1}\phi_{i_1}(x_1)\phi_{j_1}(x_1)dx_1\right)\\ &\quad\ldots\left(\sum_{j_d=i_d-1}^{i_d+1} r_n^{d,j_d}\int_{\Omega_d}\phi_{i_d}(x_d)\phi_{j_d}(x_d)dx_d\right)\end{aligned}$$

The mass matrix $\mathcal{M}$ defined below is a symmetric and tridiagonal matrix because the supports of the functions ϕ_i and ϕ_j are disjoint when $|i-j|>1$. The expressions of the terms of this mass matrix can be determined analytically.

$$(\mathcal{M})_{i,i} = 2\int_{ih}^{(i+1)h}\frac{(x-ih)^2}{h^2} = \frac{2h}{3},$$

and

$$(\mathcal{M})_{i,i+1} = \int_{ih}^{(i+1)h} \frac{(x-ih)((i+1)h-x)}{h^2} dx = \frac{h}{6}$$

Nevertheless, we can also apply a mass lumping approach to reduce the execution time when solving the linear system of equations obtained using the tridiagonal mass matrix $\mathcal{M}$. The mass lumping technique uses the trapezoidal rule for integration that gives the following approximation of the integral:

$$\int_{ih}^{(i+1)h} f(x)dx = h\frac{(f(ih)+f((i+1)h))}{2}$$

Using this formula, we can replace the matrix $\mathcal{M}$ by $\mathcal{N}$ defined as follows:

$$\mathcal{N}_{i,i} = \frac{h\left(\phi_i((i-1)h)+\phi_i(ih)\right)}{2} + \frac{h\left(\phi_i((i+1)h)+\phi_i(ih)\right)}{2} = h,$$

and

$$\mathcal{N}_{i,i+1} = \mathcal{N}_{i-1,i} = 0$$

Therefore, the matrix $\mathcal{N}$ is such that $\mathcal{N} = hI$ where the matrix I is the identity matrix. Figure 3.10 shows the convergence curves obtained using the matrices $\mathcal{M}$ and $\mathcal{N}$ respectively in the system of equations (2.19). In Figure (3.10), we can also note that solving the system (2.19) with the approximation using the mass matrix $\mathcal{M}$ is more precise than the mass lumping technique.

In Table 3.2, we compare the computational time needed to obtain 30 iterations of the PGD algorithm (2.17) to approximate a basket put option payoff function with and without mass lumping.

Dimension	Execution time in seconds with mass lumping	Execution time in seconds without mass lumping
2	1	2
3	2	7
4	2	18
5	3	34
6	4	63
7	6	97
8	11	167
9	22	298
10	80	563

Table 3.2. Execution times according to the dimension with or without mass lumping in the approximation of a basket put payoff

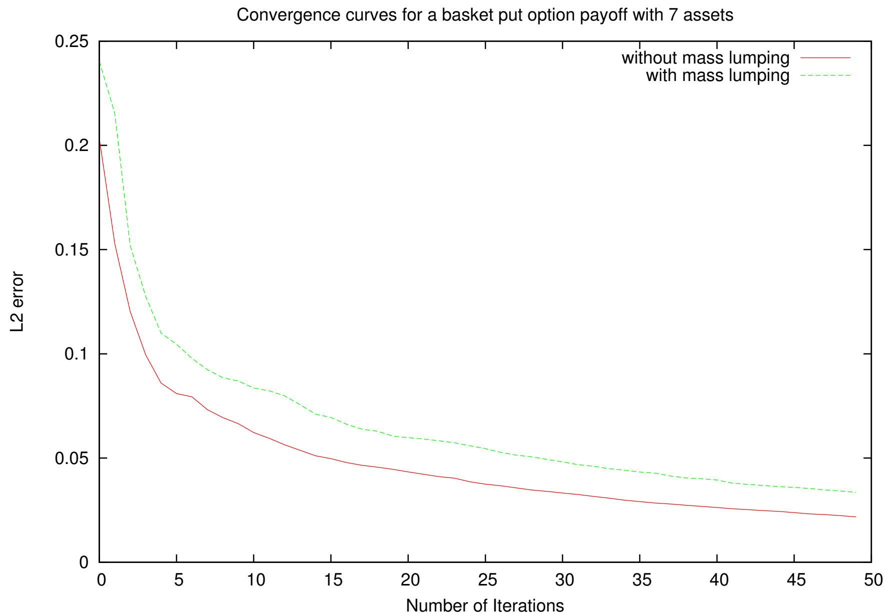


Fig. 3.10. Comparison between the convergence curves of a 7 asset put payoff obtained using the method with and without mass lumping

In Table 3.2, we consider only 11 points of discretization per dimension. As the number of points of discretization increases ($N > 21$ points), the advantage of the mass lumping method is more obvious as we can see in Table (3.3).

Number of discretization points	Execution time in seconds with mass lumping	Execution time in seconds without mass lumping
21	6	45
31	13	96

Table 3.3. Execution times according to the number of discretization points when using or not mass lumping in the approximation of a basket put payoff with 5 assets

The advantage in terms of the execution time of the mass lumping method for numerical integration is interesting because it allows to obtain relatively fast results in high-dimensional frameworks. Nevertheless, when we look for more accurate approximations we consider the approximation given by the mass matrix $\mathcal{M}$. The numerical experiments presented in this work do not consider mass lumping approximation except if it is indicated.

3.8 Pricing of a basket put using the separated approximation of the payoff

As an example to show that the approximation by a sum of tensor products makes sense, we can use this approximation as a method to obtain prices of options in the Black-Scholes framework. The price of a European option in the Black-Scholes model (see Section 4.1.1) is given by

$$\begin{aligned} P_t &= \mathbb{E}\left[e^{-r(T-t)} f(S_1(T), \ldots, S_d(T)) | \mathcal{F}_t\right] \\ &= \int_{\Omega_1 \times \ldots \times \Omega_d} e^{-r(T-t)} f(y_1, \ldots, y_d) g(T, y_1, \ldots, y_d | t, S_1(t), \ldots, S_d(t)) dy_1 \ldots dy_d \end{aligned} \tag{3.7}$$

where $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration generated by the d assets prices $S_i(t)$ $(i = 1, \ldots, d)$, f is the payoff of the option, $g(T, .|t, S_1(t), \ldots, S_d(t))$ is the joint density of the variables $S_1(T), \ldots, S_d(T)$ given the values $(S_1(t), \ldots, S_d(t))$ of the underlying assets at time t. This joint density is a log-normal law and thus has an explicit analytical expression.

Using the greedy algorithm as we saw in the previous section, we can obtain a separable approximation of the product

$$f(y_1, \ldots, y_d) g(T, y_1, \ldots, y_d | t, S_1(t), \ldots, S_d(t)),$$

and thus the integral (3.7) can be calculated very efficiently using the Fubini's rule.

In Figure 3.11, we apply this idea for the case of a basket put option on seven assets and in Table 3.4 we show results for at-the-money basket put options according to the dimension comparing the price of the method presented in this section with a Monte Carlo method with 10^4 and 10^6 simulations.

Dimension	CI 10^4	CI 10^6	Price method
3	7.54 - 7.71	7.59 - 7.61	7.59
4	8.33 - 8.54	8.48 - 8.50	8.47
5	8.35 - 8.55	8.49 - 8.51	8.48
6	8.79 - 8.99	8.88 - 8.90	8.87
7	8.94 - 9.15	9.03 - 9.05	9.00

Table 3.4. At-the money price for a basket put option in dimension d = 3, 4, 5, 6, 7. We consider the Black-Scholes model with a correlation matrix of the form $\rho_{ij} = 0.3$ for all extra-diagonal terms. The today spot is given by the d-first elements of the vector [50, 30, 50, 40, 20, 20, 50] and the volatility vector σ is obtained by the same way on [0.3, 0.1, 0.3, 0.2, 0.1, 0.2, 0.1].

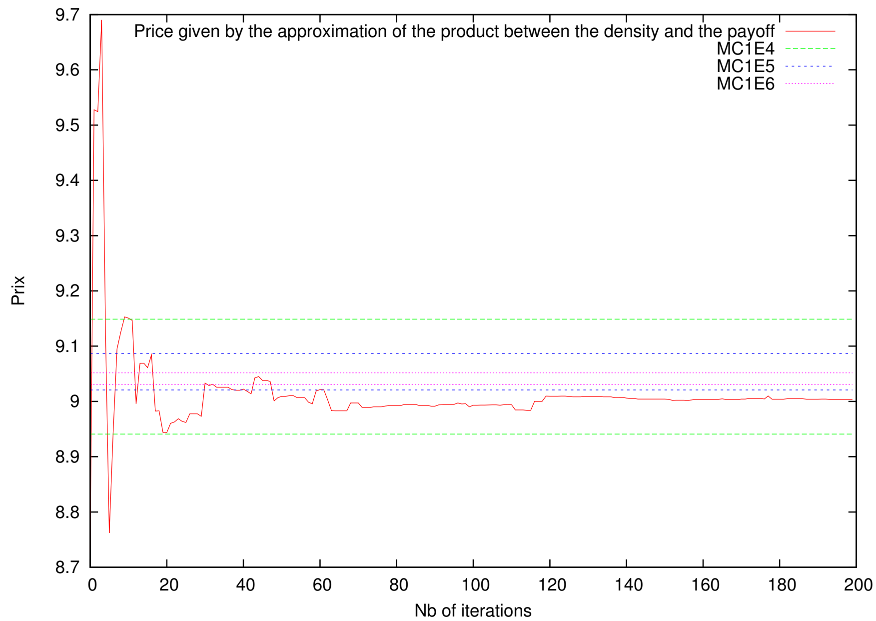


Fig. 3.11. Price of put basket option with 7 assets. The continuous curve gives the price of this financial product with respect to the number of iterations of the algorithm. The horizontal lines represent the confidence interval obtained with a Monte Carlo method using respectively 10^4, 10^5 and 10^6 iterations.

4

Application in Finance of a nonlinear approximation method for solving high-dimensional partial differential equations

Now, we will apply the greedy algorithm (2.7) presented in Section 2.2 to solve the Black-Scholes equation (1.27) and obtain the price of a European option.

4.1 The Proper Generalized Decomposition applied to the Black-Scholes partial differential equation

4.1.1 Weak formulation of the Black-Scholes partial differential equation

The Black-Scholes model in a d-dimensional framework (see Section 1.3.3) describes the dynamics of d risky assets that satisfy the following stochastic differential equations:

$$\frac{dS_i(t)}{S_i(t)} = rdt + \sigma_i dB_i(t) \text{ for all } i = 1, \ldots, d, \tag{4.1}$$

with

$$d\langle B_i, B_j \rangle_t = \rho_{ij} dt. \tag{4.2}$$

The number ρ_{ij} is the correlation between the Brownian motions B_i and B_j that drive the dynamics of the assets S_i and S_j respectively.

The coefficient σ_i represents the volatility of the asset S_i at time t and r is the risk-free instantaneous interest rate. To simplify, we assume that r and σ_i for $i = 1, \ldots d$ are constant during the period $[0, T]$. We note that the PGD method that we are proposing can be used when the risk-free interest rate is a continuous function of time and the volatility is a continuous function of time and of the asset under standard regularity assumptions (See Chapter 2 in [AP05]).

We recall that the price of a European option with payoff f and maturity T is given by the following formula

$$\mathbb{E}\left[e^{r(T-t)} f(S_1(T), \ldots, S_d(T))|\mathcal{F}_t\right].$$

Using the Markovianity of the process $(S_1(t), \ldots, S_d(t))$, it can be written as:

$$\mathbb{E}\left[e^{r(T-t)} f(S_1(T), \ldots, S_d(T))|\mathcal{F}_t\right] = P(t, S_1(t), \ldots S_d(t)), \tag{4.3}$$

where P is a deterministic function.

The function $P(t, S_1, \ldots, S_d)$ satisfies the Black-Scholes PDE which can be obtained using the Feynman-Kac theorem. The Black-Scholes equation in a d-dimensional framework is a parabolic PDE that has the following form:

$$\begin{cases} \frac{\partial P}{\partial t} + \mathcal{L}P = 0, \quad t < T, (S_1, \ldots, S_d) \in \mathbb{R}_+^d, \\ P(T, S_1, \ldots, S_d) = f(S_1, \ldots, S_d), \quad (S_1, \ldots, S_d) \in \mathbb{R}_+^d, \end{cases} \tag{4.4}$$

where the operator $\mathcal{L}$ is given by

$$\mathcal{L}P = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 P}{\partial S_i \partial S_j} \rho_{ij} \sigma_i \sigma_j S_i S_j + \sum_{i=1}^{d} r S_i \frac{\partial P}{\partial S_i} - rP.$$

Let us recall the standard framework for problem (4.4). Setting $\tau := T - t$, the time to maturity, we get the following forward parabolic problem for $\hat{P}(\tau, S_1 \ldots, S_d) = P(t, S_1 \ldots, S_d)$

$$\begin{cases} \frac{\partial \hat{P}}{\partial \tau} - \mathcal{L}\hat{P} = 0, \quad 0 < \tau \leq T, (S_1, \ldots, S_d) \in \mathbb{R}_+^d \\ \hat{P}(0, S_1, \ldots, S_d) = f(S_1, \ldots, S_d), \quad (S_1, \ldots, S_d) \in \mathbb{R}_+^d. \end{cases} \tag{4.5}$$

We note that it is possible to write the diffusion term in the operator $\mathcal{L} + r$ in a divergence form as follows:

$$\mathcal{L}\hat{P} + r\hat{P} = \frac{1}{2} \sum_{i=1}^{d} \frac{\partial}{\partial S_i} \left(\sum_{j=1}^{d} \rho_{i,j} \sigma_i \sigma_j S_i S_j \frac{\partial \hat{P}}{\partial S_j} \right) + \sum_{j=1}^{d} \left(r S_j - \frac{1}{2} \sum_{i=1}^{d} \frac{\partial}{\partial S_i} (\rho_{i,j} \sigma_i \sigma_j S_i S_j) \right) \frac{\partial \hat{P}}{\partial S_j}.$$

Therefore, if we multiply $-\mathcal{L}\hat{P}$ by a test function Q and then we integrate on $\mathbb{R}_+^d$, we obtain the following bilinear form:

$$
\begin{aligned}
b_t(\hat{P}, Q) = {} & \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \int_{\mathbb{R}_+^d} \rho_{i,j} \sigma_i \sigma_j S_i S_j \frac{\partial \hat{P}}{\partial S_j} \frac{\partial Q}{\partial S_i} \\
& - \sum_{j=1}^{d} \int_{\mathbb{R}_+^d} \left(r S_j - \frac{1}{2} \sum_{i=1}^{d} \frac{\partial}{\partial S_i} (\rho_{i,j} \sigma_i \sigma_j S_i S_j) \right) \frac{\partial \hat{P}}{\partial S_j} Q + r \int_{\mathbb{R}_+^d} \hat{P} Q.
\end{aligned}
\tag{4.6}
$$

Let us now introduce the Hilbert space

$$
\mathcal{V}(\mathbb{R}_+^d) = \left\{ v : v \in L^2(\mathbb{R}_+^d), S_i \frac{\partial v}{\partial S_i} \in L^2(\mathbb{R}_+^d), i = 1, \ldots, d \right\}
$$

and its norm

$$
\|v\|_{\mathcal{V}} = \left(\|v\|^2_{L^2(\mathbb{R}_+^d)} + \sum_{i=1}^{d} \left\| S_i \frac{\partial v}{\partial S_i} \right\|^2_{L^2(\mathbb{R}_+^d)} \right)^{\frac{1}{2}}.
$$

We have the following result for problem (4.5) (see Theorem 2.11 in [AP05]).

Theorem 4.1. *Let us assume that the matrix Ξ defined by $\Xi_{i,j} = \rho_{i,j}\sigma_i\sigma_j$ is positive-definite. Then for all $f \in L^2(\mathbb{R}_+^d)$, there exists a unique function $\hat{P} \in L^2(0,T;\mathcal{V}) \cap \mathcal{C}^0([0,T]; L^2(\mathbb{R}_+^d))$, with $\frac{\partial \hat{P}}{\partial t} \in L^2(0,T;\mathcal{V}')$ such that, for any function $\phi \in \mathcal{D}(0,T)$, for all $v \in \mathcal{V}$,*

$$
- \int_0^T \phi'(t) \left(\int_{\mathbb{R}_+^d} \hat{P}(t) v \right) dt + \int_0^T \phi(t) b_t(\hat{P}(t), v) dt = 0 \tag{4.7}
$$

and

$$
\hat{P}(t=0) = f. \tag{4.8}
$$

This result shows the existence and uniqueness of a weak solution for the problem (4.5).

In this work, our goal is to obtain the curve of prices for a put basket option which has a square-integrable payoff. The price of call basket options can be obtained by the well-known put-call parity. So, as initial condition we consider the payoff function:

$$
f(S_1, \ldots, S_d) = \left(K - \frac{1}{d} \sum_{i=1}^{d} S_i(0) \right)_+ \tag{4.9}
$$

where the constant K is the strike of the option.

Three new difficulties appear when we want to apply the PGD techniques (2.7) to solve the problem (4.5) when we compare it with the application of the PGD method to the case of the Poisson problem as we presented in Section 2.6:

1. It is a problem posed on an infinite domain.
2. It is a time-dependent problem.
3. We cannot simply recast the weak formulation (4.7) of the problem (4.5) as a minimization problem because the bilinear form (4.6) is non-symmetric.

4.1.2 Formulation on a bounded domain

The financial assets S_i for $i = 1, \dots, d$ take values in $[0, \infty)$. Consequently, we have to deal with an infinite domain. Let us then introduce the following transformations:

$$\Psi : \mathbb{R}_+ \mapsto [0,1], \quad s \mapsto \frac{s}{s + \frac{K}{d}}, \tag{4.10}$$

$$\Psi^{-1} : [0,1] \mapsto \mathbb{R}_+, \quad x \mapsto \frac{xK}{d(1-x)}. \tag{4.11}$$

As remarked by Pommier in [Pom08], the change of variables (4.10) maps bijectively $\mathbb{R}_+$ to the interval $[0, 1]$ and appears to be efficient in practice since it leads to a refined mesh around the singularity line of the payoff function. In [Pom08], Pommier explains that if we set a classical localized boundary-domain then the volume next to this singularity decays exponentially with the dimension. This change of variables allows us not to impose artificial boundary conditions contrary to classical truncation techniques. Proposition 4.1 below shows that with the change of variables (4.10), we get a well-posed problem in a bounded domain without boundary conditions.

Applying the change of variable (4.10) into the equation (4.4), we obtain:

$$\begin{cases} -\frac{\partial u}{\partial t} + \tilde{\mathcal{L}} u = 0, \\ u(0, x_1, \dots, x_d) = (K - \frac{K}{d} \sum_{i=1}^{d} \frac{x_i}{1-x_i}), \end{cases} \tag{4.12}$$

where $u(t, x_1, \dots, x_d) = P(t, S_1, \dots, S_d)$ with $S_i = \Psi(x_i)$,$(x_1, \dots, x_d) \in \Omega = (0,1)^d$ and

$$\tilde{\mathcal{L}} u = \operatorname{div}(A \nabla u) + \sum_{i=1}^{d} \left[r + \sigma_i^2 x_i - \sigma_i^2 + \frac{\sigma_i}{2} \sum_{j=1, j\neq i}^{d} \rho_{i,j} \sigma_j (2x_j - 1) \right] x_i (1 - x_i) \frac{\partial u}{\partial x_i} - ru,$$

with the matrix A given by

$$A_{i,j}(x_1,\ldots,x_d) := \frac{\rho_{i,j}\sigma_i\sigma_j}{2}x_ix_j(1-x_i)(1-x_j). \tag{4.13}$$

Then let us introduce the following Hilbert space

$$\tilde{\mathcal{V}}(\Omega) = \left\{v \in L^2(\Omega) \mid \forall\, 1 \le i \le d,\ (1-x_i)x_i\frac{\partial v}{\partial x_i} \in L^2(\Omega)\right\}, \tag{4.14}$$

endowed with the norm

$$\|v\|_{\tilde{\mathcal{V}}} = \left(\|v\|^2_{L^2(\Omega)} + |v|^2_{\tilde{\mathcal{V}}}\right)^{\frac{1}{2}} \tag{4.15}$$

where

$$|v|^2_{\tilde{\mathcal{V}}} = \sum_{i=1}^{d}\left\|x_i(1-x_i)\frac{\partial v}{\partial x_i}\right\|^2_{L^2(\Omega)}. \tag{4.16}$$

In what follows, we need the following lemma:

Lemma 4.1. *The space $\mathcal{C}_c^\infty(\Omega)$ is dense in $\tilde{\mathcal{V}}(\Omega)$.*

This lemma can be directly deduced from Lemma 2.6 in [AP05].

Corollary 4.1. *The following integration by parts formula holds:*

$$\int_\Omega \mathrm{div}(A\nabla u)v = -\int_\Omega (A\nabla u)\nabla v, \quad \forall u, v \in \tilde{\mathcal{V}}(\Omega). \tag{4.17}$$

Proof. It follows from Lemma 4.1.

Therefore, multiplying $-\tilde{\mathcal{L}}u$ by a test function $v \in \mathcal{C}_c^\infty(\Omega)$ and then using (4.17), we get the following bilinear form:

$$\tilde{b}_t(u,v) = \int_\Omega (A\nabla u)\nabla v - \int_\Omega (a\nabla u)v + \int_\Omega ruv. \tag{4.18}$$

where $a = (a_1,\ldots,a_d) : \Omega \mapsto \mathbb{R}^d$ is the vector field with i-th component given by

$$a_i(x_1,\dots,x_d) = x_i(1-x_i)\left[r+\sigma_i^2 x_i - \sigma_i^2 + \frac{\sigma_i}{2}\sum_{j=1, j\neq i}^{d} \rho_{ij}\sigma_j(2x_j-1)\right], \tag{4.19}$$

The following result holds for the the bilinear form defined in (4.18)

Lemma 4.2. *The bilinear form $\tilde{b}_t$ is continuous from $\tilde{\mathcal{V}} \times \tilde{\mathcal{V}}$, that is, there exists a constant $\overline{c}$ that does not depend on t such that for all functions $v, w \in \tilde{\mathcal{V}}$*

$$\tilde{b}_t(v,w) \leq \overline{c}\|v\|_{\tilde{\mathcal{V}}}|w|_{\tilde{\mathcal{V}}} \tag{4.20}$$

Moreover, the bilinear form $\tilde{b}_t$ verifies a Garding inequality, that is, there exist two positive constants $\underline{c} > 0$ and $\lambda > 0$ such that for all functions $v \in \tilde{\mathcal{V}}$

$$b(v,v) \geq \underline{c}|v|^2_{\tilde{\mathcal{V}}} - \lambda\|v\|^2_{L^2(\Omega)} \tag{4.21}$$

Proof. The Garding inequality is obtained by observing that the first order term satisfies the following:

$$\left|\int_\Omega p(x_1,\dots,x_d)\frac{\partial v}{\partial x_i}v\right| \leq \frac{1}{2}\left|\int_\Omega \frac{\partial p}{\partial x_i}(x_1,\dots,x_d)v^2\right|,$$

where $p(x_1,\dots,x_d)$ is a polynomial. The proof of the continuity uses the same arguments.

Thus, we obtain the following result concerning the existence and uniqueness of the solution for the weak formulation associated to the problem (4.12).

Proposition 4.1. *For all function $g \in L^2(\Omega)$, there exists a unique $u \in L^2(0,T;\tilde{\mathcal{V}}) \cap \mathcal{C}^0([0,T];L^2(\Omega))$, with $\frac{\partial u}{\partial t} \in L^2(0,T;\tilde{\mathcal{V}}')$ such that for any function $\phi \in \mathcal{D}(0,T)$, for all function $v \in \tilde{\mathcal{V}}$,*

$$-\int_0^T \phi'(t)\int_\Omega u(t)v dt + \int_0^T \phi(t)\tilde{b}(u,v)dt = 0, \tag{4.22}$$

and

$$u(t=0) = g \tag{4.23}$$

Moreover, u solution of (4.22) is related to $\hat{P}$ solution of (4.7) by the functions defined in (4.10) and (4.11).

This proposition can be deduced from Lemma 4.2 using standard techniques, see [AP05].

4.1.3 The IMEX scheme and the Black-Scholes equation as a minimization problem

To apply the PGD method (2.7), our goal is to rewrite the problem (4.12) as a minimization problem. As a first step, we propose to use an Euler scheme to discretize the problem in time. Let us consider a time discretization grid of $M+1$ points, $\tau_0 = 0 \leq \ldots \leq \tau_M = T$, where $\tau_i = i\Delta t$ and $\Delta t = \frac{T}{M}$. We introduce a time discretization of the variational formulation (4.22) where we treat explicitly the non-symmetric part of $\tilde{b}_t$ and implicitly its symmetric terms (IMEX scheme).

For $i = 1, \ldots, M$, find $u^i \in \tilde{\mathcal{V}}$ such that

$$\begin{aligned}
&\int_\Omega u^i v + \frac{\Delta t}{1+r\Delta t}\int_\Omega (A\nabla u^i)\nabla v - \frac{\Delta t}{2(1+r\Delta t)}\left[\int_\Omega (a\nabla u^i)v + \int_\Omega (a\nabla v)u^i\right] \\
&= \frac{1}{1+r\Delta t}\int_\Omega u^{i-1}v + \frac{\Delta t}{2(1+r\Delta t)}\left[\int_\Omega (a\nabla u^{i-1})v - \int_\Omega (a\nabla v)u^{i-1}\right], \quad \forall v \in \tilde{\mathcal{V}}.
\end{aligned} \tag{4.24}$$

Thus, using that the left hand side of the equation (4.24) is symmetric in u^i and v, we are led to solve the following sequence of minimization problems

For $i = 1, \ldots, M$:

$$\text{Find } u^i \in \tilde{\mathcal{V}}(\Omega) \text{ such that } u^i = \underset{u \in \tilde{\mathcal{V}}(\Omega)}{\operatorname{argmin}}\, \mathcal{E}_i(u) \tag{4.25}$$

where

$$\begin{aligned}
\mathcal{E}_i(u) = &\frac{1}{2}\int_\Omega |u|^2 + \frac{\Delta t}{2(1+r\Delta t)}\left[\int_\Omega (A\nabla u)\nabla u - \int_\Omega (a\nabla u)u\right] \\
&- \frac{1}{1+r\Delta t}\int_\Omega u^{i-1}u - \frac{\Delta t}{2(1+r\Delta t)}\left[\int_\Omega (a\nabla u^{i-1})u - \int_\Omega (a\nabla u)u^{i-1}\right].
\end{aligned} \tag{4.26}$$

Let us introduce the bilinear symmetric form $\hat{a}(u,v)$

$$\hat{a}(u,v) = \int_\Omega uv + \frac{\Delta t}{1+r\Delta t}\int_\Omega (A\nabla u)\nabla v - \frac{\Delta t}{2(1+r\Delta t)}\left[\int_\Omega (a\nabla u)v + \int_\Omega (a\nabla v)u\right], \quad \forall u, v \in \tilde{\mathcal{V}} \tag{4.27}$$

and the linear form

$$L_{i-1}(v) = \frac{1}{1+r\Delta t}\int_\Omega u^{i-1}v + \frac{\Delta t}{2(1+r\Delta t)}\left[\int_\Omega (a\nabla u^{i-1})v - \int_\Omega (a\nabla v)u^{i-1}\right], \quad \forall v \in \tilde{\mathcal{V}}. \tag{4.28}$$

We have that $\mathbb{E}_i(u) = \frac{1}{2}\hat{a}(u,u) - L_{i-1}(u)$.

4.1.4 Stability analysis for the IMEX scheme

In this section, we study the L^2-stability of the IMEX scheme (4.24). Let us consider u^n the solution of the problem (4.25) at time τ_n. In our context, the meaning of stability is given in the following definition.

Definition 4.1. *The numerical scheme* (4.24) *is L^2-stable if there exists a constant $C > 0$ that does not depend on the discretization parameter Δt such that for any initial condition u^0 and for all $n \geq 0$*

$$\|u^n\|_{L^2} \leq C\|u^0\|_{L^2}$$

We will see that the scheme (4.24) is L^2-stable under a condition on the time step Δt. For the sake of simplicity we take $r = 0$. Let us begin with the following lemma:

Lemma 4.3. *Let us assume that the matrix Ξ defined by $\Xi_{i,j} = \rho_{i,j}\sigma_i\sigma_j$, $i, j = 1, \ldots, d$, is positive-definite. Then there exists a constant $\alpha > 0$ such that*

$$\int_\Omega (A\nabla u)\, \nabla u \geq \alpha |u|^2_{\tilde{\mathcal{V}}}, \quad \forall u \in \tilde{\mathcal{V}}. \tag{4.29}$$

Proof. Using that the matrix Ξ is positive-definite, we have

$$\int_\Omega \sum_{i,j=1}^{d} \frac{\rho_{ij}\sigma_i\sigma_j}{2} x_i(1-x_i)\frac{\partial u}{\partial x_i} x_j(1-x_j)\frac{\partial u}{x_j} = \int_\Omega Y^T \Xi Y \geq \left(\min_{\lambda \in Sp(\Xi)} \lambda\right) |u|^2_{\tilde{\mathcal{V}}}$$

where $Sp(\Xi)$ denotes the set of eigenvalues of the matrix Ξ and Y is the vector such that $Y_i = x_i(1-x_i)\frac{\partial u}{\partial x_i}$. This proves (4.29) with $\alpha = \min_{\lambda \in Sp(\Xi)} \lambda$.

Now, we can state the following proposition:

Proposition 4.2. *The scheme proposed in* (4.24) *is L^2-stable under the following condition on the time step Δt*

$$\Delta t < \frac{1}{2\left(\frac{4(\|\tilde{a}\|_\infty + \|\operatorname{div}(a)\|_\infty)}{\alpha} + \frac{\alpha}{2}\right)} \tag{4.30}$$

where the constant α is defined in Lemma 4.3 and $\tilde{a}$ is the vector such that for all $i = 1, \ldots d$

$$\tilde{a}_i = \left[r + \sigma_i^2 x_i - \sigma_i^2 + \frac{\sigma_i}{2} \sum_{j=1, j\neq i}^{d} \rho_{i,j}\sigma_j(2x_j - 1) \right].$$

Proof. Let us take $r = 0$ and $v = u^i$ in the variational formulation (4.24). Thus, we obtain

$$\begin{aligned}
&\frac{1}{2\Delta t}\left(\int_\Omega |u^i|^2 - |u^{i-1}|^2\right) + \frac{1}{2\Delta t}\int_\Omega |u^i - u^{i-1}|^2 + \int_\Omega (A\nabla u^i)\nabla u^i - \int_\Omega (a\nabla u^i)u^i \\
&= \frac{1}{2}\left[\int_\Omega (a\nabla u^{i-1})u^i - \int_\Omega (a\nabla u^i)u^{i-1}\right] \\
&\Leftrightarrow \frac{1}{2\Delta t}\left(\int_\Omega |u^i|^2 - |u^{i-1}|^2\right) + \frac{1}{2\Delta t}\int_\Omega |u^i - u^{i-1}|^2 + \int_\Omega (A\nabla u^i)\nabla u^i \\
&-\frac{1}{2}\int_\Omega (a\nabla(u^i + u^{i-1}))u^i - \frac{1}{2}\int_\Omega (a\nabla u^i)(u^i - u^{i-1}) = 0.
\end{aligned} \tag{4.31}$$

Moreover, we note that $a_i = x_i(1 - x_i)\tilde{a}_i$ and that $\tilde{a} \in L^\infty(\Omega)$.

Thus, for all $\epsilon > 0$ we have that

$$\left|\int_\Omega (a\nabla u^i)(u^i - u^{i-1})\right| \leq \epsilon |u^i|^2_{\tilde{\mathcal{V}}} + \frac{\|\tilde{a}\|_\infty}{4\epsilon}\int_\Omega |u^i - u^{i-1}|^2. \tag{4.32}$$

Besides, using the integration by parts given by (4.17) to study the term $\int_\Omega (a\nabla(u^i + u^{i-1}))u^i$, we observe that

$$\begin{aligned}
\left|\int_\Omega a\nabla u^i(u^i + u^{i-1})\right| &= \left|\int_\Omega a\nabla u^i\left(2u^i + (u^{i-1} - u^i)\right)\right|, \\
&\leq \epsilon |u^i|^2_{\tilde{\mathcal{V}}} + \frac{\|\tilde{a}\|_\infty}{4\epsilon}\left(\int_\Omega 2|u^i - u^{i-1}|^2 + \int_\Omega 8|u^i|^2\right)
\end{aligned} \tag{4.33}$$

and

$$\left|\int_\Omega \operatorname{div}(a)u^i(u^i + u^{i-1})\right| \leq \frac{\|\operatorname{div}(a)\|_\infty}{4\epsilon}\left(\int_\Omega 2|u^i - u^{i-1}|^2 + \int_\Omega 8|u^i|^2\right) + \epsilon\int_\Omega |u^i|^2. \tag{4.34}$$

Then, using (4.32), (4.33) and (4.34), we deduce from (4.31) that

$$\begin{aligned}
&\left[\frac{1}{2\Delta t} - \frac{\|\tilde{a}\|_\infty}{\epsilon} - \frac{\|\operatorname{div}(a)\|_\infty}{\epsilon} - \frac{\epsilon}{2}\right]\int_\Omega |u^i|^2 + (\alpha - \epsilon)|u^i|^2_{\tilde{\mathcal{V}}} \\
&+ \left[\frac{1}{2\Delta t} - \frac{\|\tilde{a}\|_\infty}{4\epsilon} - \frac{\|\operatorname{div}(a)\|_\infty}{4\epsilon}\right]\int_\Omega |u^i - u^{i-1}|^2 \leq \frac{1}{2\Delta t}\int_\Omega |u^{i-1}|^2.
\end{aligned} \tag{4.35}$$

So, if we assume the condition (4.30) and we choose $\epsilon = \frac{\alpha}{2}$, then we can deduce the following three inequalities:

$$\alpha > \epsilon,\ \frac{1}{2\Delta t} - \frac{\|\tilde{a}\|_\infty + \|\operatorname{div}(a)\|_\infty}{\epsilon} - \frac{\epsilon}{2} > 0 \text{ and } \frac{1}{2\Delta t} > \frac{\|\tilde{a}\|_\infty + \|\operatorname{div}(a)\|_\infty}{4\epsilon}. \tag{4.36}$$

Consequently, we get

$$\begin{aligned}
\int_\Omega |u^i|^2 &\leq \frac{1}{1 - C\Delta t} \int_\Omega |u^{i-1}|^2 \\
&\leq (1 + 2C\Delta t) \int_\Omega |u^{i-1}|^2 \\
&\leq (1 + 2C\Delta t)^M \int_\Omega |u^0|^2 \\
&\leq e^{2CT} \int_\Omega |u^0|^2,
\end{aligned}$$

where $C = 2\left(\frac{(\|\tilde{a}\|_\infty + \|\operatorname{div}(a)\|_\infty)}{\epsilon} + \frac{\epsilon}{2}\right)$ is a constant which is independent of the discretization parameter Δt.

4.1.5 Implementation of the Proper Generalized Decomposition techniques for the Black-Scholes partial differential equation

To simplify the notation we consider the case of only three dimensions, but the definition of the algorithm and all the equations below can be easily generalized to a d-dimensional framework.

We recall that the PGD method will generate the function u^i in the separated representation:

$$u^i(x_1, x_2, x_3) = \sum_{k \geq 1} r_k^i \otimes s_k^i \otimes t_k^i(x_1, x_2, x_3).$$

The PGD algorithm (4.25) is defined as follows: For $i = 1, \dots, M$, iterate on $n \geq 1$

$$(r_n^i, s_n^i, t_n^i) \in \underset{\substack{r \in \tilde{\mathcal{V}}(\Omega_1), \\ s \in \tilde{\mathcal{V}}(\Omega_2), \\ t \in \tilde{\mathcal{V}}(\Omega_3)}}{\operatorname{argmin}} \frac{1}{2}\hat{a}(r \otimes s \otimes t, r \otimes s \otimes t) - L_{i-1}(r \otimes s \otimes t) - \hat{a}\left(\sum_{k=1}^{n-1} r_k^i \otimes s_k^i \otimes t_k^i, r \otimes s \otimes t\right) \tag{4.37}$$

where $\hat{a}$ is defined by (4.27) and L_{i-1} by (4.28).

Then, the Euler equation associated with the problem (4.37), that is used in practice to implement the algorithm, is given by:

Find $(r_n^i, s_n^i, t_n^i) \in \tilde{\mathcal{V}}(\Omega_1) \times \tilde{\mathcal{V}}(\Omega_2) \times \tilde{\mathcal{V}}(\Omega_3)$ such that for any functions $(r, s, t) \in \tilde{\mathcal{V}}(\Omega_1) \times \tilde{\mathcal{V}}(\Omega_2) \times \tilde{\mathcal{V}}(\Omega_3)$

$$\begin{aligned} &\hat{a}(r_n^i \otimes s_n^i \otimes t_n^i, r \otimes s_n^i \otimes t_n^i + r_n^i \otimes s \otimes t_n^i + r_n^i \otimes s_n^i \otimes t) = L_{i-1}(r \otimes s_n^i \otimes t_n^i + r_n^i \otimes s \otimes t_n^i + r_n^i \otimes s_n^i \otimes t) \\ &+\hat{a}\left(r \otimes s_n^i \otimes t_n^i + r_n^i \otimes s \otimes t_n^i + r_n^i \otimes s_n^i \otimes t, \sum_{k=1}^{n-1} r_k^i \otimes s_k^i \otimes t_k^i\right) \end{aligned} \tag{4.38}$$

Henceforth, we will consider without loss of generality $s = 0$ and $t = 0$ in order to study in detail each term of this Euler equation (4.38). We recall that the Euler equation is solved using a fixed point procedure as in (2.20).

Remark 4.1. *All the high-dimensional integrals in* (4.38) *are easily calculated using Fubini's rule because the functions in these integrals are separable except for* $i = 1$ *where the term* u^0 *appears as follows:*

$$\int_{\Omega_1 \times \Omega_2 \times \Omega_3} r \otimes s_n^i \otimes t_n^i(x_1, x_2, x_3) u^0(x_1, x_2, x_3) dx_1, dx_2 dx_3, \tag{4.39}$$

The idea used to overcome this practical obstacle is to approximate, in a preliminary step, the initial condition u^0 *of the Black-Scholes PDE by a sum of tensor products as explained in Section 3.1. Once the initial condition* u^0 *has a separated approximation:*

$$u^0(x_1, x_2, x_3) = \sum_{k \geq 1} r_k^0 \otimes s_k^0 \otimes t_k^0(x_1, x_2, x_3).$$

the integral (4.39) *is easy to compute using Fubini's rule.*

Using the space discretization described in Section 3.1 and the notation given by (3.1), the following vectors will be used:

$$\mathbf{r_n^i} = [r_{n,0}^i, \ldots, r_{n,N}^i]^T, \quad \mathbf{s_n^i} = [s_{n,0}^i, \ldots, s_{n,N}^i]^T, \quad \mathbf{t_n^i} = [t_{n,0}^i, \ldots, t_{n,N}^i]^T,$$

Given the fact that all the terms in the equation (4.38) admits a separated representation, the equation (4.38) can be written in a matrix form.

The following matricial expressions allow us to deduce the matricial form for the equation (4.38):

$$\int_{\Omega_1 \times \Omega_2 \times \Omega_3} (r_n^i \otimes s_n^i \otimes t_n^i)(r \otimes s_n^i \otimes t_n^i) = [\mathbf{t_n^i}^T M \mathbf{t_n^i}][\mathbf{s_n^i}^T M \mathbf{s_n^i}] M \mathbf{r_n^i},$$

$$\int_{\Omega_1\times\Omega_2\times\Omega_3} \left(A\nabla(r_n^i\otimes s_n^i\otimes t_n^i)\right)\nabla(r\otimes s_n^i\otimes t_n^i) = \left(\frac{\sigma_2^2}{2}[\mathbf{t_n^i}^T M\mathbf{t_n^i}][\mathbf{s_n^i}^T L\mathbf{s_n^i}] + \frac{\sigma_3^2}{2}[\mathbf{t_n^i}^T L\mathbf{t_n^i}][\mathbf{s_n^i}^T M\mathbf{s_n^i}]\right) M\mathbf{r_n^i}$$
$$+\left(\frac{\rho_{1,2}\sigma_1\sigma_2}{2}[\mathbf{t_n^i}^T M\mathbf{t_n^i}][\mathbf{s_n^i}^T D\mathbf{s_n^i}] + \frac{\rho_{1,3}\sigma_1\sigma_3}{2}[\mathbf{t_n^i}^T D\mathbf{t_n^i}][\mathbf{s_n^i}^T M\mathbf{s_n^i}]\right)$$
$$(D+D^T)\mathbf{r_n^i} + \frac{\sigma_1^2}{2}[\mathbf{s_n^i}^T M\mathbf{s_n^i}][\mathbf{t_n^i}^T M\mathbf{t_n^i}]L\mathbf{r_n^i},$$

$$\int_{\Omega_1\times\Omega_2\times\Omega_3} \left(a\nabla(r_n^i\otimes s_n^i\otimes t_n^i)\right)(r\otimes s_n^i\otimes t_n^i) = [\mathbf{s_n^i}^T M\mathbf{s_n^i}][\mathbf{t_n^i}^T M\mathbf{t_n^i}]B\mathbf{r_n^i} + \left(\frac{\rho_{1,2}\sigma_1\sigma_2}{2}[\mathbf{s_n^i}^T C\mathbf{s_n^i}][\mathbf{t_n^i}^T M\mathbf{t_n^i}]\right.$$
$$\left.+\frac{\rho_{1,3}\sigma_1\sigma_3}{2}[\mathbf{t_n^i}^T C\mathbf{t_n^i}][\mathbf{s_n^i}^T M\mathbf{s_n^i}]\right) D\mathbf{r_n^i} + \left(\frac{\rho_{1,2}\sigma_1\sigma_2}{2}[\mathbf{t_n^i}^T M\mathbf{t_n^i}][\mathbf{s_n^i}^T D\mathbf{s_n^i}]\right.$$
$$\left.+\frac{\rho_{1,3}\sigma_1\sigma_3}{2}[\mathbf{t_n^i}^T D\mathbf{t_n^i}][\mathbf{s_n^i}^T M\mathbf{s_n^i}]\right) C\mathbf{r_n^i} + \left([\mathbf{t_n^i}^T M\mathbf{t_n^i}][\mathbf{s_n^i}^T B\mathbf{s_n^i}]\right.$$
$$+\frac{\rho_{2,3}\sigma_2\sigma_3}{2}[\mathbf{t_n^i}^T C\mathbf{t_n^i}][\mathbf{s_n^i}^T D\mathbf{s_n^i}] + [\mathbf{t_n^i}^T B\mathbf{t_n^i}][\mathbf{s_n^i}^T M\mathbf{s_n^i}]$$
$$\left.+\frac{\rho_{2,3}\sigma_2\sigma_3}{2}[\mathbf{t_n^i}^T D\mathbf{t_n^i}][\mathbf{s_n^i}^T C\mathbf{s_n^i}]\right) M\mathbf{r_n^i}$$

where the matrices M, L, B, C, D are explicitly computable tridiagonal matrices of size $N\times N$, with N the number of intervals in each direction. The computation of the components for these matrices boils down to one-dimensional integrals.

In this way, solving (4.38) with a fixed point procedure allows us to obtain for a fixed i such that $1\le i\le N$, the n-th term of the sum $\sum_{k=1}^n r_k^i\otimes s_k^i\otimes t_k^i$ which is an approximation of the solution at time $t_i = i\Delta t$ of the problem (4.12).

4.2 Numerical results

4.2.1 Testing the method against an analytical solution

In this part, we apply the PGD algorithm defined in (4.37) to solve the problem (4.12) with the following initial condition:

$$u(0,x_1,\dots,x_d) = \left(K - \prod_{i=1}^d \frac{x_i}{(1-x_i)}\right)^+ \tag{4.40}$$

for which the solution is analytically known.

Using the Feynman-Kac theorem (4.3) we get that the solution of the PDE (4.5) is given by $\mathbb{E}\left[e^{-rT}(K-\prod_{i=1}^d S_T^i)^+\right]$, which is possible to calculate analytically in the Black-Scholes model. We have:

$$\mathbb{E}\left[e^{-rT}(K-\prod_{i=1}^d S_T^i)^+\right] = e^{-rT}K\mathbb{P}\left(K>\prod_{i=1}^d S_T^i\right) - e^{-rT}\mathbb{E}\left[\prod_{i=1}^d S_T^i \mathbf{1}_{\left\{K>\prod_{i=1}^d S_T^i\right\}}\right] \tag{4.41}$$

We get the quantity $\mathbb{P}\left(K > \prod_{i=1}^d S_T^i\right)$ as follows:

$$\mathbb{P}\left(K > \prod_{i=1}^d S_T^i\right) = \mathbb{P}\left(e^{\sum_{i=1}^d X_T^i} < \frac{K}{\prod_{i=1}^d S_0^i}\right)$$
$$= \mathbb{P}\left(Y < \log\left(\frac{K}{\prod_{i=1}^d S_0^i}\right)\right)$$

where $X_T^i = (r - \frac{\sigma_i^2}{2})T + \sigma_i W_T^i$ and $Y = \sum_{i=1}^d X_T^i$ is a normal random variable with mean equal to $\sum_{i=1}^d \left(r - \frac{\sigma_i^2}{2}\right) T$ and variance given by $\sum_{i=1}^d \sum_{j=1}^d \rho_{ij}\sigma_i\sigma_j T$.

Besides, we remark that:

$$\mathbb{E}\left[\prod_{i=1}^d S_T^i \mathbf{1}_{\left\{K > \prod_{i=1}^d S_T^i\right\}}\right] = \mathbb{E}\left[e^Y \mathbf{1}_{\{e^Y < \frac{K}{\prod_{i=1}^d S_0^i}\}}\right] \prod_{i=1}^d S_0^i$$

where the last term can be calculated analytically. We remark that the analytic solution (4.41) is not separable with respect to each coordinate.

We present in figure 4.1 a numerical example of the solution obtained with our algorithm and the analytic solution. In figure 4.2 we see the same surface but intersected with the plane $x_1 = x_2$.

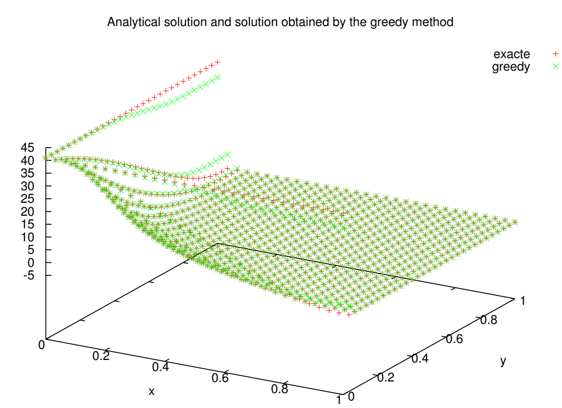


Fig. 4.1. The analytical solution and the numerical one obtained with our method for the problem (4.5) with the initial condition (4.40) in a two-dimensional framework. For this example, we consider $\Delta t = \frac{1}{100}$ and $\Delta x = \frac{1}{30}$.

Figure 4.3 shows the convergence curves, i.e, the L^2 relative error with respect to the number of iterations of the algorithm according to the dimension. We note that the number of iterations needed to obtain convergence increases as the dimension of the problem increases.

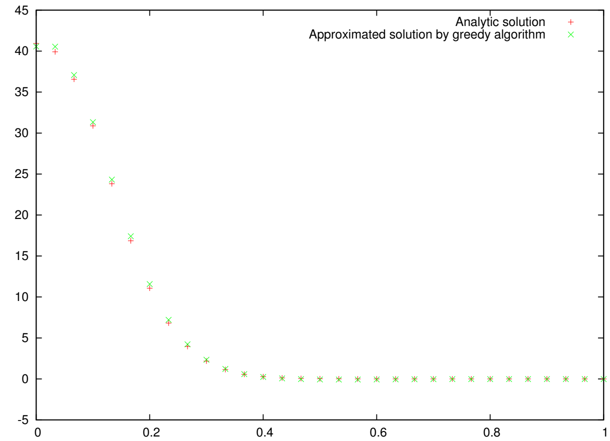


Fig. 4.2. The analytical solution and the numerical one obtained with our method for the problem (4.5) with the initial condition (4.40) in a two-dimensional framework. We represent the intersection between the surface in figure 4.1 and the plane $x_1 = x_2$. For this example, we consider $\Delta t = \frac{1}{100}$ and $\Delta x = \frac{1}{30}$.

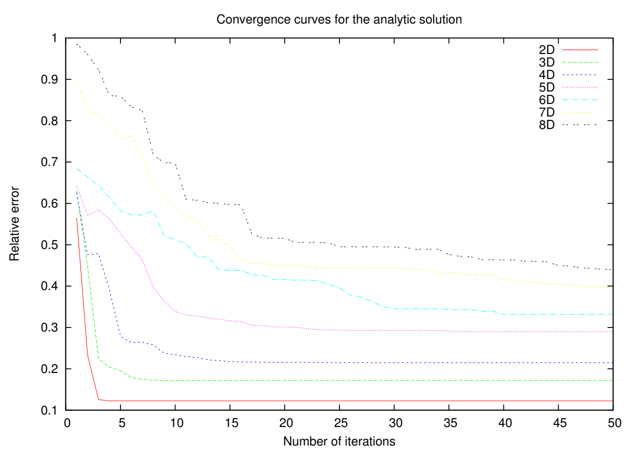


Fig. 4.3. Convergence curves for the solution at time T of the equation (4.5) with initial condition given by (4.40). To obtain this curves, we consider $\Delta x = 0.1$ for each dimension and $\Delta t = 0.01$.

4.2.2 Results on the Black-Scholes equation

In this section, we show the results that we obtained applying the PGD method described in the previous section to the Black-Scholes equation.

Figure 4.4 represents the approximation of the solution at time T to the problem (4.5) with initial condition (4.9) obtained using the PGD approximation defined by the equations (4.38)

Figure 4.5 represent the price of a basket put option when all the assets take the same value, i.e. $S_1 = \ldots = S_d$. Precisely, Figure 4.5 compares prices obtained with different discretizations Δx.

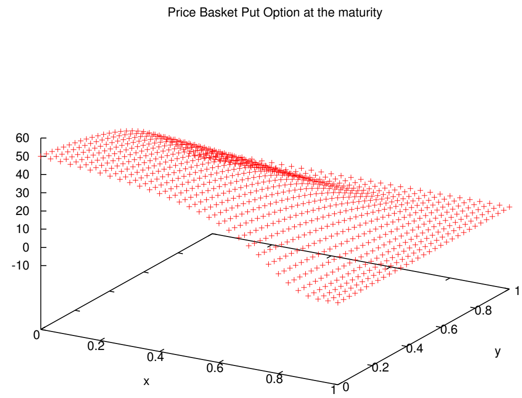


Fig. 4.4. The approximated solution obtained with our method for the problem (4.5) with the initial condition (4.9).

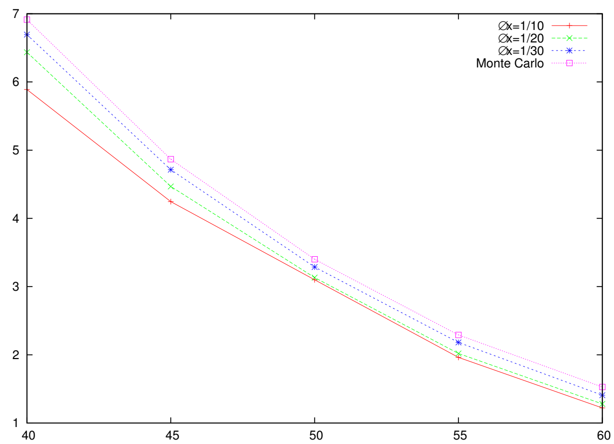


Fig. 4.5. The approximated solution obtained with our method for the problem (4.5) with the initial condition (4.9) in a four-dimensional framework. For these results, we set $\Delta x = 0.1, 0.2$ and $\Delta t = 0.01$.

In terms of computational time, the PGD approach (4.37) is not competitive compared to Monte Carlo methods when one is interested in the price of only one value of the spot, but on the other side the curve of prices is obtained for any time $t \in [0, T]$ and any price spot.

4.2.3 Application as a variance reduction method

In this part, we show that we can use the solution obtained by the PGD method described above in order to find a control variable to reduce the variance when calculating the price an of option.

We can re-write the equations (4.1) and (4.2) which define the Black-Scholes model as follows:

$$\frac{dS_t^i}{S_t^i} = rdt + \sigma^i \sum_{j=1}^{d} H_{i,j} dW_t^j \tag{4.42}$$

where W is a d-dimensional Brownian motion and the matrix H verifies $HH^t = \Sigma$ where Σ is a $d \times d$ matrix such that $\Sigma_{i,j} = 1$ if $j = i$ and $\Sigma_{i,j} = \rho_{ij}$ otherwise.

We recall that the price of a basket put option is given by

$$\mathbb{E}\left[e^{-rT} f(S_T^1, S_T^2, \ldots, S_T^d)\right]$$

where $f(S_T^1, S_T^2, \ldots, S_T^d) = \left(K - \frac{1}{d}\sum_{i=1}^{d} S_T^i\right)_+$.

Now, for the sake of simplicity, let us consider $r = 0$. Generalization to $r \neq 0$ is straightforward. Let us introduce the Kolmogorov equation:

$$\begin{cases} \partial_t \hat{P} - \frac{1}{2} A : \nabla^2 \hat{P} = 0 \\ \hat{P}(0, x) = f(x) \end{cases} \tag{4.43}$$

where $A = FH(FH)^T$ and F is a diagonal matrix such that $F_{i,i} = \sigma^i S^i$ for $i = 1, \ldots, d$. Notice that

$$\hat{P}(T, S_0) = \mathbb{E}\left[f(S_T^1, S_T^2, \ldots, S_T^d)\right].$$

Therefore, we have

$$\hat{P}(0, S_T) - \hat{P}(T, S_0) = \int_0^T FH \; \nabla\hat{P}(T-t, S_t) dB_t$$

and thus,

$$\hat{P}(T, S_0) = f(S_T) - \int_0^T FH \; \nabla\hat{P}(T-t, S_t) dB_t \tag{4.44}$$

The random variable $Y = \int_0^T FH \;\; \nabla\hat{P}(T-t, S_t) dB_t$ has zero mean and is a perfect control variable since

$$\mathrm{Var}\left[f(S_T) - Y\right] = 0.$$

As we do not know the solution $\hat{P}$, in practice, we calculate an approximation $\hat{P}^\star$ of $\hat{P}$ using the PGD algorithm presented in Section 4.1.5. Therefore, we obtain an approximated control variable

$Y^\star = \int_0^T FH \, \nabla \hat{P}^\star(T-t, S_t) dB_t$ and we can compute an approximation of $\hat{P}(T, S_0)$ by Monte Carlo, computing the following quantity:

$$\mathbb{E}\left[f(S_T) - \int_0^T FH \, \nabla \hat{P}^\star(T-t, S_t) dB_t\right].$$

We remark that this idea can be applied to any payoff function and that for a new value of S_0 we use the same approximation $\hat{P}^\star$.

In Table 4.1 and 4.2 we present the performance of our variance reduction method compared with the variance obtained with the classical method, i.e. calculating directly $\mathbb{E}[f(S_T)]$.

Dimension	Without variance reduction	With variance reduction
4	0.1233	0.0012
5	0.1204	0.0034
6	0.1197	0.0078
7	0.1245	0.0113
8	0.1257	0.0254

Table 4.1. Variance with a correlation parameter $\rho_{i,j} = 0.9$ constant between all the assets.

Dimension	Without variance reduction	With variance reduction
4	0.1256	0.0023
5	0.1248	0.0045
6	0.1230	0.0096
7	0.1199	0.0158
8	0.1232	0.0296

Table 4.2. Variance with a correlation parameter $\rho_{i,j} = 0.1$ constant between all the assets.

For two typical values of the correlation, we observe that the reduction of the variance is important, for example, up to a factor 6 in dimension 8.

4.3 Appendix: Formulas for the matrices used to solve the Black-Scholes partial differential equation

In this section, we detail the calculations of the terms appearing in the equation (4.38). As we already remarked, the multi-dimensional integrals are in fact the multiplication of 1-dimensional integrals.

Let us now analyze some terms that appear in the left hand side of the equation (4.38). Thus we have that:

$$\begin{aligned}
\int_{\Omega_1\times\Omega_2\times\Omega_3} (A\nabla(r_n^i\otimes s_n^i\times t_n^i))\nabla(r\otimes s_n^i\otimes t_n^i) =& \frac{1}{2}\sigma_1^2\int_{\Omega_3}|t_n^i|^2\int_{\Omega_2}|s_n^i|^2\int_{\Omega_1}x_1^2(1-x_1)^2(\partial_{x_1}r_n^i)(x_1)\partial_{x_1}(r(x_1))dx_1\\
&+\frac{\rho_{12}\sigma_1\sigma_2}{2}\int_{\Omega_3}|t_n^i|^2\int_{\Omega_2}x_2(1-x_2)s_n^i(x_2)\partial_{x_2}(s_n^i)(x_2)dx_2\int_{\Omega_1}x_1(1-x_1)\partial_{x_1}(r_n^i)(x_1)rdx_1\\
&+\frac{\rho_{13}\sigma_1\sigma_3}{2}\int_{\Omega_2}|s_n^i|^2\int_{\Omega_3}x_3(1-x_3)t_n^i(x_3)\partial_{x_3}(t_n^i)(x_3)dx_3\int_{\Omega_1}x_1(1-x_1)\partial_{x_1}(r_n^i)(x_1)rdx_1\\
&+\frac{1}{2}\sigma_2^2\int_{\Omega_3}|t_n^i|^2\int_{\Omega_2}x_2^2(1-x_2)^2\partial_{x_2}(s_n^i)(x_2)\partial_{x_2}(s_n^i)(x_2)dx_2\int_{\Omega_1}r_n^i(x_1)r(x_1)dx_1\\
&+\frac{\rho_{12}\sigma_1\sigma_2}{2}\int_{\Omega_3}|t_n^i|^2\int_{\Omega_2}x_2(1-x_2)s_n^i(x_2)\partial_{x_2}(s_n^i)(x_2)dx_2\int_{\Omega_1}x_1(1-x_1)r_n^i\partial_{x_1}(r)(x_1)dx_1\\
&+\frac{1}{2}\sigma_3^2\int_{\Omega_2}|s_n^i|^2\int_{\Omega_3}x_3^2(1-x_3)^2\partial_{x_3}(t_n^i)(x_3)\partial_{x_3}(t_n^i)(x_3)dx_3\int_{\Omega_1}r_n^i(x_1)r(x_1)dx_1\\
&+\frac{\rho_{13}\sigma_1\sigma_3}{2}\int_{\Omega_2}|s_n^i|^2\int_{\Omega_3}x_3(1-x_3)t_n^i(x_3)\partial_{x_3}(t_n^i)(x_3)dx_3\int_{\Omega_1}x_1(1-x_1)r_n^i\partial_{x_1}(r)(x_1)dx_1.
\end{aligned} \tag{4.45}$$

We recall that we consider a discretization in space given by $\Delta x = \frac{1}{N}$ and that is the same for each dimension. So, in order to calculate these last terms, we define the next matrices:

$$L_{i,j} = \int_\Omega x^2(1-x)^2\partial_x(\phi_i)(x)\partial_x(\phi_j)(x)dx \tag{4.46}$$

$$D_{i,j} = \int_\Omega x(1-x)\phi_i(x)\partial_x(\phi_j)(x)dx \tag{4.47}$$

The matrix L is a tridiagonal matrix of size $N+1\times N+1$ given by:

$$L_{i,j} = \begin{cases}
\frac{1}{5}(\Delta x)^3 - \frac{1}{2}(\Delta x)^2 + \frac{1}{3}(\Delta x), & \text{if } \ i=j=0\\
-\frac{1}{5}(\Delta x)^3 + \frac{1}{2}(\Delta x)^2 - \frac{1}{3}(\Delta x), & \text{if } \ i=0, j=1\\
\frac{1}{30}(\Delta x)(10-15(\Delta x)+6(\Delta x)^2), & \text{if } \ i=j=N\\
-\frac{1}{30}(\Delta x)(10-15(\Delta x)+6(\Delta x)^2), & \text{if } \ i=N, j=N-1\\
\frac{2}{15}(\Delta x)(15(\Delta x)^2 i^4 + 30(\Delta x)^2 i^2 + 3(\Delta x)^2 - 30(\Delta x)i^3 - 30(\Delta x)i + 15i^2 + 5), & \text{if } \ j=i\\
\begin{aligned}&-\frac{1}{30}(\Delta x)\left[30(\Delta x)^2 i^4 - 60(\Delta x)^2 i^3 + 60(\Delta x)^2 i^2 - 30(\Delta x)^2 i + 6(\Delta x)^2\right.\\ &\left.-60(\Delta x)i^3 + 90(\Delta x)i^2 - 60(\Delta x)i + 15(\Delta x) + 30i^2 - 30i + 10\right],\end{aligned} & \text{if } \ j=i-1\\
\begin{aligned}&-\frac{1}{30}(\Delta x)\left[30(\Delta x)^2 i^4 + 60(\Delta x)^2 i^3 + 60(\Delta x)^2 i^2 + 30(\Delta x)^2 i + 6(\Delta x)^2 - 60(\Delta x)i^3\right.\\ &\left.-90(\Delta x)i^2 - 60(\Delta x)i - 15(\Delta x) + 30i^2 + 30i + 10\right],\end{aligned} & \text{if } \ j=i+1
\end{cases}$$

The matrix D is also tridiagonal of size $N+1 \times N+1$ and it is such that:

$$D_{i,j} = \begin{cases} \frac{1}{12}(\Delta x)((\Delta x)-2), & \text{if } \begin{array}{l} i=j=0 \\ \text{or } i=N, j=N-1 \end{array} \\ -\frac{1}{12}(\Delta x)((\Delta x)-2), & \text{if } \begin{array}{l} i=j=N \\ \text{or } i=0, j=1 \end{array} \\ \frac{1}{3}(\Delta x)(2i(\Delta x)-1), & \text{if } j=i \\ \frac{1}{12}(\Delta x)(6i^2(\Delta x)-4i(\Delta x)+(\Delta x)-6i+2), & \text{if } j=i-1 \\ -\frac{1}{12}(\Delta x)(6i^2(\Delta x)+4i(\Delta x)+(\Delta x)-6i-2), & \text{if } j=i+1 \end{cases}$$

Another term that appear in the left hand side of the equation (4.38) is the next one

$$\begin{aligned} \int_{\Omega_1\times\Omega_2\times\Omega_3} (a\nabla(r_n^i\otimes s_n^i\otimes t_n^i))(r\otimes s_n^i\otimes t_n^i) = \\ = \int_{\Omega_3}|t_n^i(x_3)|^2dx_3\int_{\Omega_2}|s_n^i(x_2)|^2dx_2\int_{\Omega_1}(r+\sigma_1^2x_1-\sigma_1^2)x_1(1-x_1)\partial_{x_1}(r_n^i)(x_1)r(x_1)dx_1 \\ +\frac{\rho_{12}\sigma_1\sigma_2}{2}\int_{\Omega_3}|t_n^i(x_3)|^2dx_3\int_{\Omega_2}(2x_2-1)|s_n^i(x_2)|^2dx_2\int_{\Omega_1}x_1(1-x_1)\partial_{x_1}(r_n^i)(x_1)r(x_1)dx_1 \\ +\frac{\rho_{13}\sigma_1\sigma_3}{2}\int_{\Omega_2}|s_n^i(x_2)|^2dx_2\int_{\Omega_3}(2x_3-1)|t_n^i(x_3)|^2dx_3\int_{\Omega_1}x_1(1-x_1)\partial_{x_1}(r_n^i)(x_1)r(x_1)dx_1 \\ +\frac{\rho_{21}\sigma_1\sigma_2}{2}\int_{\Omega_3}|t_n^i(x_3)|^2dx_3\int_{\Omega_2}x_2(1-x_2)\partial_{x_2}(s_n^i)(x_2)s_n^i(x_2)dx_2\int_{\Omega_1}(2x_1-1)r_n^i(x_1)r(x_1)dx_1 \\ +\int_{\Omega_3}|t_n^i(x_3)|^2dx_3\int_{\Omega_2}(r+\sigma_2^2x_2-\sigma_2^2)x_2(1-x_2)\partial_{x_2}(s_n^i)(x_2)s_n^i(x_2)dx_2\int_{\Omega_1}r_n^i(x_1)r(x_1)dx_1 \\ +\frac{\rho_{23}\sigma_2\sigma_3}{2}\int_{\Omega_3}(2x_3-1)|t_n^i(x_3)|^2dx_3\int_{\Omega_2}x_2(1-x_2)\partial_{x_2}(s_n^i)(x_2)s_n^i(x_2)dx_2\int_{\Omega_1}r_n^i(x_1)r(x_1)dx_1 \\ +\frac{\rho_{31}\sigma_1\sigma_3}{2}\int_{\Omega_2}|s_n^i(x_2)|^2dx_2\int_{\Omega_3}x_3(1-x_3)\partial_{x_3}(t_n^i)(x_3)t_n^i(x_3)dx_3\int_{\Omega_1}(2x_1-1)r_n^i(x_1)r(x_1)dx_1 \\ +\int_{\Omega_2}|s_n^i(x_2)|^2dx_2\int_{\Omega_3}(r+\sigma_3^2x_3-\sigma_3^2)x_3(1-x_3)\partial_{x_3}(t_n^i)(x_3)t_n^i(x_3)dx_3\int_{\Omega_1}r_n^i(x_1)r(x_1)dx_1 \\ \frac{\rho_{23}\sigma_2\sigma_3}{2}\int_{\Omega_2}(2x_2-1)|s_n^i(x_2)|^2dx_2\int_{\Omega_3}x_3(1-x_3)\partial_{x_3}(t_n^i)(x_3)t_n^i(x_3)dx_3\int_{\Omega_1}r_n^i(x_1)r(x_1)dx_1 \end{aligned} \tag{4.48}$$

Therefore, we define the next matrices:

$$B_{ij} = \int_\Omega (r+\sigma^2x-\sigma^2)x(1-x)\partial_x(\phi_i)(x)\phi_j(x)dx \tag{4.49}$$

and

$$C_{ij} = \int_{\Omega} (2x-1)\phi_i(x)\phi_j(x)dx \tag{4.50}$$

B is a tridiagonal matrix of size $N+1 \times N+1$ such that:

$$B_{i,j} = \begin{cases} \frac{1}{60}(\Delta x)(3v^2(\Delta x)^2 + 5(\Delta x)r - 10v^2(\Delta x) - 10r + 10v^2), & \text{if } \ i=j=0 \\ -\frac{1}{60}(\Delta x)(3v^2(\Delta x)^2 + 5(\Delta x)r - 10v^2(\Delta x) - 10r + 10v^2), & \text{if } \ i=0, j=1 \\ -\frac{1}{60}(\Delta x)(5v^2(\Delta x) - 3v^2(\Delta x)^2 - 10r + 5(\Delta x)r), & \text{if } \ i=j=N \\ \frac{1}{60}(\Delta x)(5v^2(\Delta x) - 3v^2(\Delta x)^2 - 10r + 5(\Delta x)r), & \text{if } \ i=N, j=N-1 \\ \frac{1}{30}(\Delta x)(10v^2 - 40v^2(\Delta x)i + 20(\Delta x)ri + 3v^2(\Delta x)^2 + 30v^2(\Delta x)^2 i^2 - 10r), & \text{if } \ j=i \\ \begin{aligned} &-\frac{1}{60}(\Delta x)\left[45v^2(\Delta x)^2 i - 20v^2 + 15(\Delta x)r - 30ri - 30v^2(\Delta x) + 80v^2(\Delta x)i - 40(\Delta x)ri \right. \\ &\left. -12v^2(\Delta x)^2 + 30v^2(\Delta x)^2 i^3 - 60v^2(\Delta x)^2 i^2 - 60v^2(\Delta x)i^2 + 30(\Delta x)ri^2 + 30v^2 i + 20r\right] \end{aligned} & \text{if } \ j=i-1 \\ \begin{aligned} &-\frac{1}{60}(\Delta x)\left[15v^2(\Delta x)^2 i + 10v^2 + 5(\Delta x)r - 30ri - 10v^2(\Delta x) - 40v^2(\Delta x)i + 20(\Delta x)ri \right. \\ &\left. +3v^2(\Delta x)^2 + 30v^2(\Delta x)^2 i^3 + 30v^2(\Delta x)^2 i^2 - 60v^2(\Delta x)i^2 + 30(\Delta x)ri^2 + 30v^2 i - 10r\right]. \end{aligned} & \text{if } \ j=i+1 \end{cases}$$

C is also a tridiagonal matrix of size $N+1 \times N+1$ such that:

$$C_{i,j} = \begin{cases} \frac{1}{6}(\Delta x)(-2 + (\Delta x)), & \text{if } \ i=j=0 \\ \frac{1}{6}(\Delta x)((\Delta x) - 1), & \text{if } \ i=0, j=1 \\ -\frac{1}{6}(\Delta x)(-2 + (\Delta x)), & \text{if } \ i=j=N \\ -\frac{1}{6}(\Delta x)((\Delta x) - 1), & \text{if } \ i=N, j=N-1 \\ \frac{2}{3}(\Delta x)(2i(\Delta x) - 1), & \text{if } \ j=i \\ \frac{1}{6}(\Delta x)(2i(\Delta x) - (\Delta x) - 1), & \text{if } \ j=i-1 \text{ or } j=i+1 \end{cases}$$

We do not develop the expression of the right hand side of the equation (4.38) because its terms can be obtained using the same matrices B, C, D, L and M.

5

Perspectives of the application of the Proper Generalized Decomposition method for option pricing

In this section, we present some perspectives on the first part of this manuscript: the problem of the American options and the use of non-continuous dynamics for the asset value.

5.1 Pricing using the characteristic function, application to Bermudan options

Before treating the pricing of an American option, it is important to remark that the PGD approach can be applied as well in the case of another type of option apart from European and American options and that can be considered as an intermediate point between them. This type of option is known as Bermudan option and the holder of one of these contracts can exercise the option only at certain days between the present time and the maturity. For example, for an option on a single asset, if we denote the payoff by the function ϕ, then the pricing function is such that

$$p(t_i^+, S) = \max\Big(p(t_i^-, S), \phi(S)\Big)$$

at each exercising time $t_i \in [0, T]$ and satisfies the PDE of Black-Scholes (4.4) between the exercising times.

Using the same idea as in the Section 3.8, we can compute numerical integration on the Fourier domain. This method can readily be applied for solving problems under various asset price dynamics, for which the characteristic function (the Fourier transform of the probability density function) is available. This is the case for models from the class of regular affine processes [DFS03] which also includes the exponentially affine jump-diffusion class of [DPS00]. Many methods based on Fourier transform techniques are used on the one dimensional asset class of models [CMS99, LO08, KL09]. As a future work, an idea would be to extend these methods to the multi-dimensional framework.

Let us introduce $\mathbb{T} = t_1, \ldots, t_M$, $t_0 \leq t_1 \leq \ldots \leq t_M = T$ a set of exercise dates where we define the time between two exercise dates as $\delta_i = t_i - t_{i-1}$. Applying the dynamic programming principle, the pricing formula for a Bermudan option is given by

$$p(t_M, x) = \phi(t_M, x),$$

$$p^+(t_{m-1}, x) = e^{-r\delta_i} \int_{\mathbb{R}^d} p(t_m; y) g(y|t_{m-1}, t_m; x) dy, \tag{5.1}$$

$$p(t_{m-1}, x) = \max\left(\phi(t_{m-1}, x), p^+(t_{m-1}, x)\right)$$

for m=M,. . . ,2 and

$$p(t_0, x) = e^{-r\delta_1} \int_{\mathbb{R}^d} p(t_1; y) g(y|t_0, t_1; x) dy$$

Here, the probability density function of $y = X_{\tilde{t}}^{x,t}$ under the risk-neutral measure is denoted $g(y|t, \tilde{t}, x)$. In this case, $y = \log\left(\frac{S_{t_m}}{K}\right)$, $p(t, x)$ is the value of the option, $p^+(t, x)$ is the continuation value and $\phi(t, x)$ is the payoff.

Considering a multi-dimensional truncated domain of integration to proceed numerically, we obtain that the payoff is L^2-integrable. Mathematically, we take the function ϕ such that

$$\phi(t_m, x) = \phi(t_m, x)\mathbf{1}_{x\in\Omega}, \quad 1 \leq m \leq M. \tag{5.2}$$

where Ω is a bounded subset of $\mathbb{R}^d$.

Using the approximation (5.2), we can compute the Fourier transform of (5.1)

$$\mathcal{F}\left(p^+(t_{m-1}, .)\right)(\xi) = e^{-r\delta_m} \int_{\mathbb{R}^2\times\mathbb{R}^2} p(y, t_m) g(y|t_{m-1}, t_m, x) e^{ix\xi} dy dx$$

Assuming that the probability density function comes from an independent increment process (for example a Lévy process), that is

$$g(y|t_{m-1}, t_m, x) = g(y - x|t_{m-1}, t_m, 0)$$

Then, it holds

$$\mathcal{F}\left(p^+(t_{m-1},.)\right)(\xi) = e^{-r\delta_m}\mathcal{F}\left(p(t_m,.)\right)(\xi)\Phi(t_{m-1},t_m;-\xi) \tag{5.3}$$

where $\Phi(t,\tilde{t},\xi) = \int_{\mathbb{R}^d} f(z|t,\tilde{t},0)e^{iz\xi}dz$.

In conclusion, we can propose the following steps to compute the price of the Bermudan option:

1. Compute an approximation of $p(t_m,.)$ as a sum of tensor products using the PGD approach (2.17) to obtain the function $\tilde{p}$ as follows:

$$\tilde{p}(.,t_m) = \sum_{n=0}^{N} r_n^1 \otimes \ldots \otimes r_n^d$$

2. Apply the Fubini's rule to get

$$\mathcal{F}\left(\tilde{p}(t_m,.)\right) = \sum_{n=0}^{N} \mathcal{F}(r_n^1) \otimes \ldots \otimes \mathcal{F}(r_n^d)$$

3. Convolve by density function using (5.3)

$$\mathcal{F}\left(\tilde{p}^+(t_{m-1},.)\right)(\xi) = e^{-r\delta_m}\mathcal{F}\left(\tilde{p}(t_m,.)\right)(\xi)\Phi(\xi).$$

4. Compute an approximation of $\mathcal{F}\left(\tilde{p}^+(t_{m-1},.)\right)$ as a sum of tensor products by the PGD method (2.17) and calculate the inverse Fourier transform using the Fubini's rule.

5.2 The problem of the American options

In this section we propose an application of the PGD method for solving the problem of the American options. This type of options can be exercised at any time up to the maturity, consequently, the price of an American option with payoff ϕ and maturity T is given by

$$p(t,x) = \sup_{\tau\in\mathcal{T}_{[t,T]}} \mathbb{E}\left[e^{-\int_t^\tau rds}\phi(S_\tau^{t,x})\right],$$

for stopping times τ between t and T, for $t \in [0,T]$ and $x \geq 0$, where $\mathcal{T}_{[t,T]}$ denotes the set of stopping times τ of the filtration $\mathcal{F}_\tau$.

It is possible to show that the pricing of an American option amounts to solving (see [AP05]):

$$\begin{cases} \max\left\{\partial_t u + Lu, u-\phi\right\} = 0 \text{ in } [0,T]\times\mathbb{R}_+^d \\ u(0,.) = \phi \text{ in } \mathbb{R}_+^d \end{cases} \tag{5.4}$$

where L is defined by (1.26) and the problem (5.4) is equivalent to the following problem, called the obstacle problem

$$\begin{cases} Lu - \dfrac{\partial u}{\partial t} \leq 0, \quad \text{in } \mathcal{R}_T \\ u \geq \phi, \quad \text{in } \mathcal{R}_T \\ (u-\phi)\left(Lu - \dfrac{\partial u}{\partial t}\right) = 0, \quad \text{in } \mathcal{R}_T \\ u(0,x) = \phi(0,x), \quad x \in \mathbb{R}^d \end{cases} \tag{5.5}$$

To apply the PGD approach to the problem (5.4), we propose to use the idea presented in [CEL12] and that allows to apply the PGD method to the obstacle problem (5.5) by penalizing the constraints. In order to use this idea we can discretize in time, obtaining:

$$\begin{cases} \max\{\frac{u^{n+1}-u^n}{\Delta t} + Lu^{n+1}, u^{n+1} - \phi\} = 0 \text{ in } [0,T] \times \mathbb{R}^d_+ \\ u(0,.) = \phi \text{ in } \mathbb{R}^d_+ \end{cases}$$

Finally, we propose to penalize the constraints:

$$\frac{u^{n+1}-u^n}{\Delta t} + L^S u^{n+1} + L^A u^n + \frac{1}{\epsilon}(u^{n+1} - \phi)_+ = 0 \tag{5.6}$$

Again, the problem in u^{n+1} is symmetric and thus can be rewritten as a minimization problem needed to apply the PGD method.

There exists theoretical results that indicate that the approach (5.6) is reasonable (see [CEL12]).

Part II

Liquidity risk, limit order book modeling and market microstructure

6

Survey on market impact models

6.1 Introduction

Financial markets are places where buyers and sellers can exchange different types of assets. One of the goals of a financial market is to allow companies to fund themselves through different types of financial products. This implies that an important part of the economy is influenced by the financial markets. The understanding of how these financial markets work is therefore very important.

As a consequence of the technological innovations and regulatory changes, the structure of the financial markets has evolved in the past years. Initiatives such as the MiFID in Europe and the Regulation National Markets System in United States have appeared to deregulate markets and then to improve the service given to the investors by stimulating the competition between the markets. The deregulation has been one possible reason for which liquidity risk and market microstructure have become a hot topic in mathematical finance in the past years.

Market microstructure is the study of how the markets work. Thus, this branch of finance is interested in understanding the details of the trading mechanisms. Among the topics treated by the market microstructure we have the price formation and price discovery, the market structure and design issues, and the information and disclosure.

The importance of the market microstructure can also be seen from the common assumption of frictionless markets, that means markets without transaction costs. This hypothesis is considered in many important models in mathematical finance such as the well-known Black-Scholes model. Because of the assumption, the price process does not depend on the trading strategy used by the agents. As the time scale becomes smaller and smaller, this assumption is less coherent.

Liquidity risk is a decisive concept in market microstructure. This risk takes into account finding the counterpart for the buy or sell order that the trader needs to execute. Liquidity is therefore a key to knowing how prices will change after the action of the agents in the markets through some measures of the liquidity of the asset such as the bid-ask spread or the market depth.

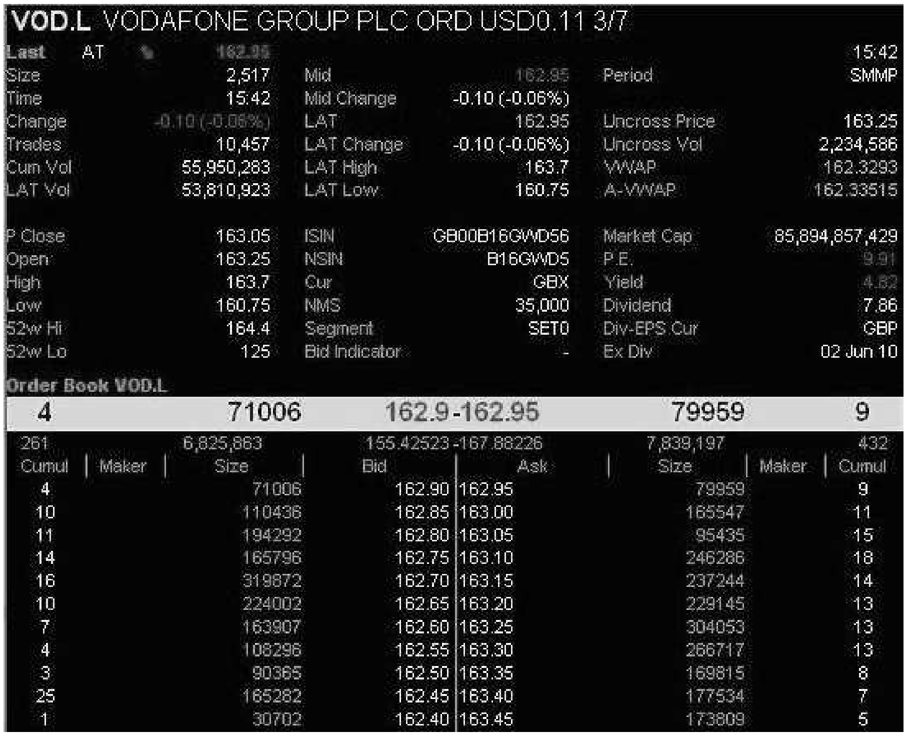
Markets are also supposed to be perfectly liquid in classical models of mathematical finance, an assumption that guarantees the perfect hedging as in the case of the Black-Scholes model for example. In practice, however, this is not true; trading large volumes of orders shifts the price of the asset in the opposite direction to the one most beneficial to the investor. As a consequence, the price impact can be seen as the difference between the realized price and the price before the order was executed.

We cannot take further the discussion on market impact models without introducing the concept of an order, which constitutes an instruction from customers to brokers to buy or sell on the markets. In the financial markets, there exist different types of orders, the more standard ones being the market orders and the limit orders. On the one hand a market order allows the client to execute his order immediately and to get the best price available on the market. On the other hand, a limit order gives to the investor the possibility of buying or selling a share at a specified price. When the execution of the order is greater than the price for the trader, he will generally use a market order. A limit order might not be able to be executed at all but it would allow the investors to have control over the price that they might pay for executing their orders.

A database of limit orders constitutes a limit order book. In other words, investors place orders to buy or sell a given quantity of stock in the limit order book. If a market order arrives in the limit order book, then it is matched with the best current price proposed on the market and if a limit order is placed in the limit order book, this will be stored waiting for a market order that would execute it. Figure 6.1 illustrates a limit order book for specific shares. Orders can only be placed on a price grid and at each moment in time, the number of waiting buy (or sell) orders is stored for each price. For a given price, orders are executed according to the FIFO rule as soon as two orders match. As we can observe, the structure of Limit Order Books (LOB) is very complex, so an exhaustive modeling of its dynamics would not lead, for example, to draw quantitative conclusions on an optimal trading strategy. One has therefore to propose models that can grasp important features of the LOB structure but that allow to find analytical results.

Quotation rules are changing and, in general, more information is now available. In particular, it is possible to know at any given time the number of awaiting orders for certain stock and to obtain a record of all past transactions. A natural question arises: how to use at the best this information to execute buy or sell orders optimally? In other words, what is the behavior of traders who want to minimize their trading cost?

Given that the size of the order can be large, the trader needs to find an optimal strategy which consists in using the resilience of the markets to help split this large order into smaller orders. This resilience of markets can be explained through the supply and demand of the assets. Hence, by splitting the large order, the trader will decrease the impact of his trade on the price of the asset. But, if the investor waits, he will increase his market risk. In other words, the market impact forces slow trading



VOD.L VODAFONE GROUP PLC ORD USD0.11 3/7

Last AT	162.95				15:42
Size	2,517	Mid	162.95	Period	SMMP
Time	15:42	Mid Change	-0.10 (-0.06%)		
Change	-0.10 (-0.06%)	LAT	162.95	Uncross Price	163.25
Trades	10,457	LAT Change	-0.10 (-0.06%)	Uncross Vol	2,234,586
Cum Vol	55,950,283	LAT High	163.7	VWAP	162.3293
LAT Vol	53,810,923	LAT Low	160.75	A-VWAP	162.33515
P Close	163.05	ISIN	GB00B16GWD56	Market Cap	85,894,857,429
Open	163.25	NSIN	B16GWD5	P.E.	9.91
High	163.7	Cur	GBX	Yield	4.82
Low	160.75	NMS	35,000	Dividend	7.86
52w Hi	164.4	Segment	SET0	Div-EPS Cur	GBP
52w Lo	125	Bid Indicator	-	Ex Div	02 Jun 10

Order Book VOD.L

Cumul	Maker	Size	Bid	Ask	Size	Maker	Cumul
4		71006	162.9	162.95	79959		9
261		6,825,863	155.42523	167.88226	7,839,197		432
4		71006	162.90	162.95	79959		9
10		110436	162.85	163.00	165547		11
11		194292	162.80	163.05	95435		15
14		165796	162.75	163.10	246286		18
16		319872	162.70	163.15	237244		14
10		224002	162.65	163.20	229145		13
7		163907	162.60	163.25	304053		13
4		108296	162.55	163.30	266717		13
3		90365	162.50	163.35	169815		8
25		165282	162.45	163.40	177534		7
1		30702	162.40	163.45	173809		5

Fig. 6.1. Limit order book of Vodafone shares containing the shares quoted on the London Stock Exchange. Source: http://www.londonstockexchangegroup.com/investor-relations/group-at-a-glance/electronic-trading/electronic-trading-main.htm

but the market risk forces quick trading. The idea is then to find a good trade-off between liquidity and the market risk.

It has been reported, for example in [AFS10], that optimal execution is a recurrent problem for practitioners because traders liquidate about twenty percent of the daily volume of shares. More dangerous practical situations can also exist as an example of application of the optimal execution, for instance, the well-known case of Jérôme Kerviel at the Société Générale. Here the bank had to liquidate his large speculative positions. Another example of this type is when a financial institution holds a large quantity of government bonds and as the default probability of that country increases, the value of these assets quickly decreases. The institution therefore would be interested in liquidating this portfolio as quickly as possible. It is natural that in selling everything at the outset is not optimal because, given the limited liquidity of the market, this financial institution would have to sell its bonds very cheaply.

It is possible to distinguish three different time scales for the trading activity. Asset management activities generate tasks like the hedging of positions that can be set in long-scale time periods, for example, days or weeks. These activities imply that it is necessary to manage positions and to execute orders on the market in a matter of minutes. The last time scale for trading is the so-called high-frequency trading which is performed in intervals of time from 10^{-6} to 1 seconds. The market impact models that we will study in this work apply to the intermediate time scale of trading, that goes typically from a minute to some hours.

In Section 6.2 we will define the optimal execution problem and we will analyze some market impact models that the literature has proposed in order to solve it. In the same section we will present the main results of these models and the connections and differences between them. We will also study the concept of price manipulation and the existence of optimal strategies for the optimal execution problem.

6.2 Market impact models

The aim of this Section is to compare the market impact models proposed by the literature and to establish links between them, emphasizing the advantages and the drawbacks of each.

6.2.1 Definition of the optimal execution problem

The problem that we will study in this manuscript is the optimal liquidation of a portfolio with $\mathbb{x}$ shares by a large trader who can place orders over a period of time $[0, T]$. We consider that $\mathbb{x} > 0$ (resp. $\mathbb{x} < 0$) corresponds to a selling (resp. buying) program. By large investor, we understand a trader who executes a large volume of orders and therefore is able to modify the price of the exchanged asset. In practice, regulators define the concept of large trader in order to supervise markets and to avoid manipulation and fraud. In the literature, these large traders are differentiated from the normal traders (also called noise traders) who, on average, do not modify the price of the asset. In this work, we will analyze the case of only one large trader. See [BP05] and [CLV07] for frameworks that consider more than one trader in competition with each other.

In order to find a solution for this optimal liquidation problem, we need to analyze the impact of the actions of the agents on the markets, that is, to model the market impact. This impact is closely related to the liquidity in financial markets described by the market impact models. We could think that market impact depends on many variables like bid-ask spread or market capitalization; nevertheless, the market impact models that we will study are related to the shift in price and to the exchanged volume.

For the sake of clarity, we would like to introduce here a general framework for studying this optimal liquidation problem. First, let us assume the existence of four price processes: $(A_t^0, t \geq 0)$, $(A_t, t \geq 0)$, $(B_t^0, t \geq 0)$ and $(B_t, t \geq 0)$ on a given filtered probability space $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P})$. The processes $(A_t^0, t \geq 0)$ and $(B_t^0, t \geq 0)$ represent respectively the best ask price and the best bid price of the asset unaffected by the large trader. The processes $(A_t, t \geq 0)$ and $(B_t, t \geq 0)$ are the actual best ask price and best bid price where the orders of the large investor are taken into account.

To simplify, most of the models proposed in the literature neglect the bid-ask spread and thus consider only two processes: $(S_t^0, t \geq 0)$ and $(S_t, t \geq 0)$ which represent respectively the price when the large trader is inactive and the price when the orders of the large investor have an influence on the limit order book. It is important to note that for liquid assets the size of the bid-ask spread is usually one tick as has been empirically shown, see Cont and De Larrard in [CD11] and the references therein. Thus, in practice, a zero bid-ask spread for liquid assets is not far from reality in the time scale where our model is set. As the unaffected price process is determined by the actions of the noisy traders, it is usually supposed to be a martingale. Other reasons for this consideration are that the time period of trading is normally short and that including a nonzero drift would lead to profitable strategies that have to be differentiated from price manipulation strategies.

Even if in practice the problem of the optimal execution of orders make sense in the discrete-time framework, it is interesting to analyze its outcomes in both the continuous and discrete time cases. For example, an important topic to study is whether the optimal strategies are stable when we increase the number of dates for trading.

For discrete-time strategies, we assume that at most $N+1$ market orders can be placed by the large trader. In this case, an admissible strategy is given by an increasing sequence $\tau_0 = 0 \leq \ldots \leq \tau_N = T$ of stopping times and random variables $\xi_0, \ldots, \xi_N$, where ξ_i is the size of the order placed at time τ_i, ($\xi_i < 0$ corresponds to a sell order, and $\xi_i > 0$ holds for a buy order) such that

- $\mathbb{x} + \sum_{i=0}^{N} \xi_i = 0$, i.e. the investor liquidates his portfolio of $\mathbb{x}$ shares,
- ξ_i is $\mathcal{F}_{\tau_i}$-measurable, and
- $\exists M \in \mathbb{R}, \forall\, 0 \leq i \leq N,\ \xi_i \geq M$, a.s.

This third hypothesis will be needed for technical reasons and can be seen as the fact that the investor cannot propose (or take out) an infinite amount of liquidity on the markets.

In addition, if we analyze continuous-time strategies, we can define an admissible strategy as a stochastic process $(X_t)_{t\geq 0}$ such that

- $X_0 = \mathbb{x}$ and $X_{T+} = 0$,
- X is $\mathcal{F}_t$-adapted and left-continuous, and
- the function $t \in [0, T+] \mapsto X_t$ has finite and a.s. bounded total variation.

This process $(X_t)_{t\geq 0}$ represents the number of shares that remain to liquidate at time t. We can observe that in the discrete-time strategy case, we have $X_t = \mathbb{x} + \sum_{i=0}^{N} \xi_i \mathbf{1}_{\tau_i < t}$ and we can deduce the assumptions on the continuous case from the hypothesis of the discrete case.

The question about the objective of the traders which determines the choice of the criterion to state the problem of optimal execution is not closed. The trader could be interested only in minimizing the expected execution cost or he could take into account the variance of his strategy as well, arguing that the expected execution cost criterion misses the volatility risk.

There exist three types of market impacts that the literature considers according to the time that they shift the price of the underlying asset: the permanent impact is a permanent change in the price such as an impact due to a new information on fundamentals of the asset; the instantaneous or temporary impact only influences the actual order; and the transient impact modifies the price after the placement of an order but decays over time. We will see that the permanent and temporary impacts are first introduced in the literature leading to a first family of models. The consideration of the transient impact comes later to produce a second type of model.

6.2.2 First family of models (immediate and permanent price impact)

In the first works mentioned in the literature for treating the optimal execution problem, the impact created by the large trader was incorporated directly into the dynamic of the asset exogenously. In these models, the price impact depends on the size of the order and on the speed of liquidation of the position.

We can differentiate these models according to the time framework: discrete-time and continuous-time models.

Discrete-time models

Among these models, Bertsimas and Lo [BL98] and Almgren and Chriss [AC00] studied first the problem of the optimal execution. Precisely, they consider the following discrete-time model

$$S_{t_n} = S^0_{t_n} + \tilde{\gamma} \sum_{i=0}^{n} \xi_i, \quad n = 0, \dots, N, \tag{6.1}$$

where $\tilde{\gamma} > 0$ and in order to take into account the impact on the price, the price process S_{t_n} consists of an arithmetic random walk and a linear function of the order size. The time schedule of trading is considered to be a uniform discretization of the period $[0, T]$, so $t_n = \frac{nT}{N}$.

Bertsimas and Lo consider in [BL98] that the aim of the large trader is to minimize the average cost of his trading program. Hence the goal is to find $\xi_0, \dots, \xi_N$ such that

$$\min_{\xi_0,\ldots,\xi_N,\text{ such that: } \mathtt{x}+\sum_{i=0}^N \xi_i=0} \mathbb{E}\left[\mathcal{C}^{\mathrm{BL}}(\xi)\right], \tag{6.2}$$

where

$$\mathcal{C}^{\mathrm{BL}}(\xi) := \sum_{i=0}^N \xi_i S_{t_i}. \tag{6.3}$$

As we already mentioned, the strategy ξ^{naive} that consists in selling the entire quantity $\mathtt{x}$ allows to eliminate the future risk on the value of the asset but its cost can be very high. Precisely, in this case, we have the following cost for this strategy

$$\mathcal{C}^{\mathrm{BL}}(\xi^{\text{naive}}) = -S_0^0\mathtt{x} + \tilde{\gamma}\mathtt{x}^2$$

and the variance of this strategy is zero.

On the other hand, the strategy that is optimal in the framework of [BL98] and that consists of a constant rate of trading given by $\xi_i^{\star,\mathrm{BL}} = -\frac{\mathtt{x}}{N}$, minimizes the expected cost given by (6.3) but its variance can be large. The cost of this strategy is given by

$$\mathbb{E}\left[\sum_{k=0}^N \frac{-\mathtt{x}}{N} S_{t_k}\right] = -\mathtt{x}S_0^0 + \frac{(N+1)\tilde{\gamma}\mathtt{x}^2}{2N}$$

and its variance is equal to

$$\mathrm{Var}\left[\sum_{k=0}^N \frac{\mathtt{x}}{N} S_{t_k}\right] = \frac{\mathtt{x}^2}{N^2}\mathbb{E}\left[\sum_{k=0}^N \left\{2(N-k)+1\right\}(S_{t_k}^0)^2\right] - \mathtt{x}^2(S_0^0)^2$$

The differences between these strategies $\xi^{\star,\mathrm{BL}}$ and ξ^{naive} illustrate the fact that for this type of model where the price impact is only permanent and temporary, the optimal strategy has to be a trade-off between the cost of the strategy and its variance.

Removing the assumption that the investor is not risk averse and based on a similar model for the price asset, in [AC00] Almgren and Chriss use a mean-variance risk criterion for the investor setting a level of risk aversion of the agent.

Thus, the goal of Almgren and Chriss in [AC00] is to find the optimal strategy for the problem

$$\mathbb{E}\left[\mathcal{C}^{\mathrm{AL}}(\xi)\right] + \lambda \mathrm{Var}\left[\mathcal{C}^{\mathrm{AL}}(\xi)\right] \tag{6.4}$$

for different values of the risk aversion parameter λ and $\mathcal{C}^{\mathrm{AL}}$ is the cost in the Almgren and Chriss framework. Indeed, they show that for each value of $\lambda \geq 0$, there exists a unique optimal strategy that minimizes the criterion (6.4).

Let us consider a uniform time grid, that means $t_i = \frac{iT}{N}$ for $i = 0, \ldots, N$ and $\delta t = t_i - t_{i-1} = \frac{T}{N}$. In [AC00], Almgren and Chriss obtain the solution for the optimal execution problem in a discrete case. Precisely, as in [BL98], the permanent and temporary impact appear directly in the dynamic of the asset

$$S_{t_k} = S_{t_{k-1}} + \sigma\sqrt{\delta t}\chi_k - \delta t g\left(\frac{\xi_k}{\delta t}\right), \tag{6.5}$$

where σ is the volatility of the asset, ξ_k the size of the order executed at time t_k and $(\chi_k)_{k=1,..,n}$ is a sequence of i.i.d. random variables with zero mean and unit variance. The function g represents the permanent impact and, in the equation (6.5), this impact is considered in the average rate of trading $\frac{\xi_k}{\delta t}$ in the interval from t_{k-1} to t_k.

In this model, the temporary impact is taken into account through the function h appearing in the price $\tilde{S}_{t_k}$ obtained by the trader in the market at the time t_k

$$\tilde{S}_{t_k} := S_{t_{k-1}} - h\left(\frac{\xi_k}{\delta t}\right).$$

Using that $\mathcal{C}^{\mathrm{AC}} := \sum_{i=0}^{N} \xi_i \tilde{S}_{t_i}$, Almgren and Chriss in [AC00] find an explicit optimal strategy in the case of a linear permanent impact function g and of a temporary impact function h given by

$$h\left(\frac{\xi_k}{\delta t}\right) = \epsilon \operatorname{sgn}(\xi_k) + \eta \frac{\xi_k}{\delta t}$$

for some positive constants ϵ, η.

For a parameter of risk aversion of the trader $\lambda > 0$, this optimal strategy is determined as follows:

$$X_k^{\star,\mathrm{AL}} = \frac{\sinh(\kappa(T - t_k))}{\sinh(\kappa T)} \mathbb{x}, \; k = 0, \ldots, N$$

where

$$\kappa \sim \sqrt{\frac{\lambda\sigma^2}{\eta}} + O(\delta t), \; \delta t \to 0.$$

In this case, this model allows an easy implementation and to obtain some insights about the parameters. We note also that the case $\lambda \to 0$ in (4.4) gives $X^\star = \mathbb{x}\frac{(T-t_j)}{T}$ which is the optimal strategy already found by Bertsimas and Lo in [BL98]. We recall that the process X_k denotes the amount of shares to detain at time t_k but not the size of the order executed by the trader.

A disadvantage of discrete-time models as [BL98, AC00] is that the time trades are fixed in advance of the trading which is clearly not realistic.

Continuous-time models

The Almgren and Chriss model is also studied in a continuous-time framework by Almgren in [Alm03]. He obtains analytic optimal strategies for a temporary impact given by a power-law function $h(\dot{X}_t) = \eta \dot{X}_t^k$ where $k > 0$ and for a linear permanent impact given by $g(\dot{X}_t) = \gamma \dot{X}_t$. We remark that $\dot{X}_t$ represents the speed of execution of the asset S_t.

In this same continuous-time framework, let us assume that the permanent impact and the temporary impact are linear functions of $X_t - \mathbb{x}$ and $\dot{X}_t$ respectively. We note that the quantity $X_t - \mathbb{x}$ represents the amount of shares already executed in the market by the large trader. This implies that the price S_t is given as follows:

$$S_t = S_t^0 + \gamma(X_t - \mathbb{x}) + \eta \dot{X}_t \tag{6.6}$$

Here, the parameter γ describes the permanent impact of the orders and η the temporary impact. We can mention that a linear market impact of the orders is not satisfactory as shows the empirical study [AHLT05] by Almgren *et al.* For more details, see [Alm03] and the references therein. We recall that the simplicity to handle the obtained expressions and to get closed formulas is the main reason why a linear impact is used in the literature.

We note that, if we sell the asset, the price has to decrease, so we impose $\gamma \geq 0, \eta \geq 0$, $\mathbb{x} > 0$ and $X_T = 0$.

Therefore, we can define the cost as follows

$$\mathcal{C}^{\mathrm{A}}(X_t) := \mathbb{E}\left[\int_0^T S_t dX_t\right], \tag{6.7}$$

that can be further written as:

$$\mathbb{E}\left[\int_0^T S_t dX_t\right] = -S_0^0\mathbb{x} + \gamma\mathbb{E}\left[\int_0^T X_t dX_t\right] - \gamma\mathbb{x}\mathbb{E}\left[\int_0^T dX_t\right] + \eta\mathbb{E}\left[\int_0^T \dot{X}_t^2 dt\right]$$
$$= -S_0^0\mathbb{x} + \frac{\gamma}{2}\mathbb{x}^2 + \eta\mathbb{E}\left[\int_0^T \dot{X}_t^2 dt\right]. \tag{6.8}$$

Thus, we can deduce that if $\eta = 0$ (i.e. no temporary impact is considered) then the cost is independent of the strategy, and any strategy is optimal. We remark that, in discrete-time, the same conclusion can be obtained.

Otherwise, when $\eta > 0$, as we minimize the cost (6.7), we can use the Jensen inequality

$$\frac{1}{T}\int_0^T \dot{X}_t^2 dt \geq \left(\frac{1}{T}\int_0^T \dot{X}_t dt\right)^2, \tag{6.9}$$

to deduce that

$$\dot{X}_t = -\frac{\mathbb{x}}{T} \tag{6.10}$$

because (6.9) is an equality when (6.10) is verified.

We observe that the first family of models presents the disadvantage that the price impact functions are deterministic functions which depend only on the quantity exchanged and do not take into account the dynamic of the limit order book, neglecting the effect of the trade over the supply and demand in the limit order book. It is intuitive that when an investor wants to place orders during a period of time, the response of the limit order book to the previous exchanges has to be one of the most important points to analyze. This is the motivation for the second family of models.

6.2.3 Second family of models (transient price impact)

Two new aspects appear in the model considered by Obizhaeva and Wang in [OW05] in order to consider the limitations of the first family of models. To be more specific, they first propose a model for the dynamic of the limit order book from which they can deduce the market impact of the large trader, instead of considering that the price impact is given as a fundamental as in [BL98, AC00]. They also introduce the concept of transient impact which means that the price shift remains after the order is placed but it decays because of the resilience of the limit order book. Hence, the idea of transient impact naturally appears when the response of the limit order book is taken into account.

In [OW05], the limit order book is represented by a constant distribution of shares to sell and to buy, respectively: the ask and bid part of this limit order book. This is called a one side model because the trader is not allowed to sell when his goal is to buy $\mathbb{x}$ shares; or conversely, he is not allowed to buy when his purpose is to sell. When a sell market order arrives, it takes away the liquidity placed at the left of the best bid and affects the price linearly with respect to the volume exchanged (the symmetric situation happens when a buy market order arrives). The limit order book will recover from the impact of the sell market order through resilience. On the other hand, when the large trader is inactive, the best bid and the best ask move according to the action of the noise traders, which add new orders in between the bid-ask spread. The price impact created by this mechanism is neither temporary nor permanent, but transient. Considering a linear permanent impact in this model is possible and it does not change the optimal strategy. Figures 6.2, 6.3, 6.4 and 6.5 illustrate the idea of the model.

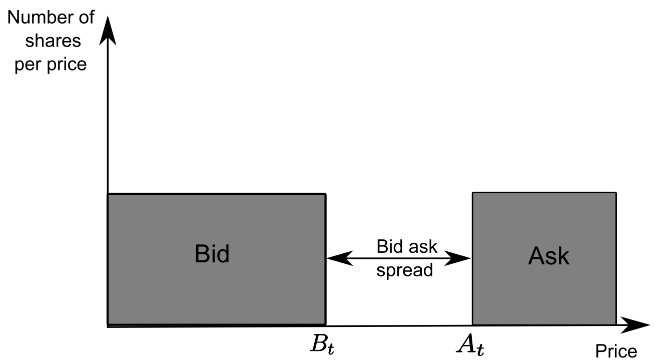


Fig. 6.2. Limit order book of an asset according to the model introduced by Obizhaeva and Wang in [OW05]. The sell and buy orders at that moment are the ask and bid sides of the book. The bid ask spread is the difference between the best ask price A_t and the best bid price B_t.

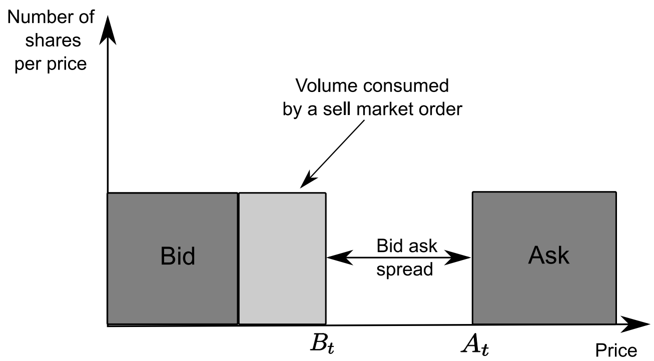


Fig. 6.3. Limit order book of an asset in the model introduced by Obizhaeva and Wang in [OW05]. A sell market order arrives to the limit order book and it is matched with the best prices offered in the bid part of the limit order book.

We can say that the resilience of the market was taken into account implicitly in the previous researches [BL98] and [AC00] when treating temporary impact with instantaneous recovery of the price after the execution of the order. In other words, in these models, the resilience is considered infinite. Meanwhile, the permanent price impact corresponds to a zero resilience case for the limit order book. We also see that the resilience reflects the fact that there are new orders appearing in

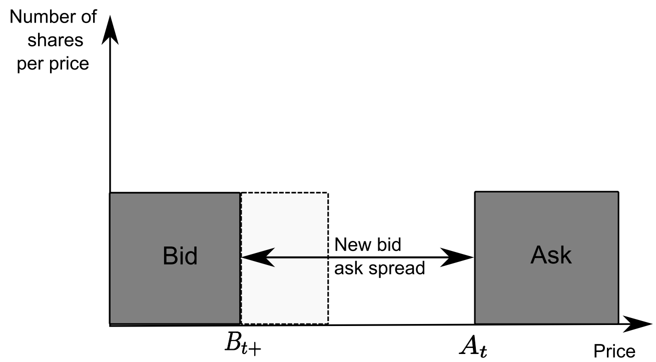


Fig. 6.4. Limit order book of an asset in the model introduced by Obizhaeva and Wang in [OW05]. B_{t+} is the new best bid price after the execution of the sell market order that arrives in Figure 6.3 and a new bid-ask spread is created.

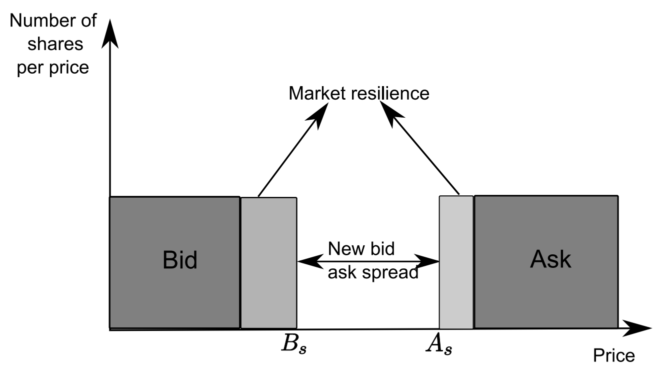


Fig. 6.5. Limit order book of an asset in the model introduced by Obizhaeva and Wang in [OW05]. When the large trader is inactive, new orders appear in between of the best ask price and the best bid price generating a new bid-ask spread.

the bid-ask spread and not in other places like the depth of the limit order book. In other words, we remark that the density of the limit order book does not change in the depth of the limit order book. We therefore deduce that the optimal strategy in the type of model proposed by [OW05] has to be a trade-off between the recovery effect on the price (that is the resilience of the limit order book) and the time constraint to liquidate the portfolio. The effect of the resilience is detected in such empirical studies as [BHS95, BGPW04, BP03] and [RW05]. Consequently, [OW05] gives new types of optimal strategies, that are combinations of discrete and continuous trades, and that are unlike the optimal strategies in [BL98] and [AC00] that consist only of continuous trades. Figure 6.6 compares the different actions of the traders according to the different types of models.

There is an important assumption in the model proposed by Obizhaeva and Wang in [OW05], that is, the quantity of shares offered on the market is constant for all prices S_t. This is what is called a block-shaped limit order book. The fact that in practice there are no block-shaped limit order books, only non-constant limit order books, has been shown in empirical studies as [BGPW04, BP03, RW05, BHS95]. The constant form of the limit order book implies that the price impacts considered in this

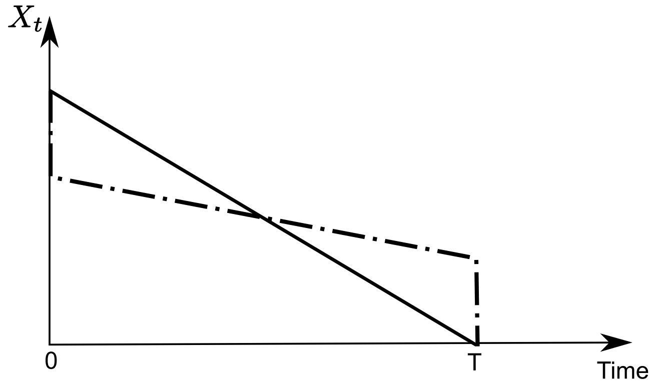


Fig. 6.6. Comparison between the optimal strategies obtained in [OW05] and [AC00]. The broken line represents the optimal strategy in the framework introduced by Obizhaeva and Wang in [OW05] and the continuous line is the optimal strategy deduced by Almgren and Chriss in [AC00].

framework are linear with respect to the volume of executed orders. However, the linearity in the impact on the traded volume is not true in reality; the impact is empirically proven to be non-linear in [Alm03, AHLT05].

Let us explain this further. If we want to estimate the market impact generated by the execution of an order of size ξ_t according to the model presented in [OW05], in other words, the change of price from S_t to S_{t+}, we can consider the following equation:

$$\int_{S_t}^{S_{t+}} q dx = \xi_t \tag{6.11}$$

given that the distribution of shares offered on the market is a constant q with respect to the price. So, (6.11) implies that the price impact given by $S_{t+} - S_t$ is linear with respect to the volume traded ξ_t because

$$S_{t+} - S_t = \frac{\xi_t}{q}.$$

We note that the optimal strategies that result from all the models discussed above [BL98, AC00, OW05] are static (or deterministic). This means that these optimal strategies can be determined in advance to the trading and that they are independent of the martingale unaffected price process. To observe this, let us consider the cost of executing a market order of size ξ_t

$$\pi_t(\xi_t) := \int_{S_t}^{S_t+} (S_t^0 + x) q dx = S_t^0 \xi_t + q \int_{S_t}^{S_t+} x dx$$

Consequently, the accumulated cost of an admissible strategy ξ is

$$\sum_{i=0}^{N} \pi_{t_i}(\xi_i) = \sum_{i=0}^{N} S^0_{t_i}\xi_i + q\sum_{i=0}^{N} \int_{S_{t_i}}^{S_{t_i+}} xf(x)dx \tag{6.12}$$

By considering the process X_t, we can re-write the first sum in the right hand of (6.12) side as follows

$$\sum_{i=0}^{N} S^0_{t_i}\xi_i = \sum_{i=0}^{N} S^0_{t_i}(X_{t_{i+1}} - X_{t_i}) = -\mathbb{x}S^0_0 - \sum_{i=1}^{N} X_{t_n}(S^0_{t_n} - S^0_{t_{n-1}}) \tag{6.13}$$

Given that the strategy ξ is admissible, the process X_{t_n} is $\mathcal{F}_{t_{n-1}}$-measurable. We remark that X_t is also a bounded process. Therefore using the martingale property of the unaffected price process S^0, the expectation of (6.13) is equal to $-S_0\mathbb{x}$. Moreover, we note that the difference $S_{t+} - S_t$ does not depend on the martingale process S^0_t and that this difference is deterministically determined when the values $\xi_0(w), \xi_1(w), \ldots, \xi_N(w)$ are given. Hence, there exists a deterministic function $C^{(i)} : \mathbb{R}^{N+1} \mapsto \mathbb{R}$ such that

$$\sum_{i=0}^{N} \int_{S_t}^{S_{t_i+}} xf(x)dx = C^{(i)}(\xi_0, \ldots, \xi_N)$$

Consequently, we obtain that

$$\mathcal{C}(\xi) = -S_0\mathbb{x} + \mathbb{E}\left[C^{(i)}(\xi_0, \ldots, \xi_N)\right],$$

which implies that the optimal strategy do not react to the changes in the price process.

We observe that the second family of models still does not consider important features of the market and of the traders' behavior. In the next section, we mention some works that take into account some of these features.

6.2.4 Other models

We remark that in all the previous models, the large investor can only place market orders, but in the markets, investors have a large variety of types of orders. See, for instance, the work of Gueant, Lehalle and Fernandez [FGL12] and Avellaneda and Stoikov [AS08] for a large trader who can use limit orders in his strategies.

A different type of modelling is proposed by Cont and De Larrard in [CD11]. Using a Markovian queueing model of the limit order book that describes essentially what is happening at the bid and ask prices, they deduce some quantities of interest such as the distribution of the duration between price changes and the probability of an upward movement in the price, conditional on the state of the order book. Nevertheless, the problem of the optimal execution is not treated.

The problem of optimal liquidation has been studied as well by Schied and Schöneborn in [SS09]. Their model is based on a continuous version of the model used in [AC00] with $T = \infty$. Considering temporary and permanent impact that are linear and using a stochastic control approach, they also analyze the adaptive liquidation strategies that react to the price changes, which are unlike the static strategies given by the previous models. Schied and Schöneborn state that, in general, these adaptive liquidation strategies give higher expected utility than the static strategies only in the case of investors who have non-constant risk aversion. These adaptive strategies were also studied by Almgren in [AL07].

We can go deeper in the explanation of the differences between the models with adaptive and static strategies by trying to interpret the goal of each type model. Thus, in [OW05] the model describes the mechanism of the limit order book dynamic from an average point of view. In practice, the limit order book is a discrete function of prices which are determined by the size of the tick. The continuous distribution of shares represented by the shape of the limit order book is then an assumption of what is happening in average on the markets. As a consequence of this fact and that the associated time scale is wider, obtaining static strategies in the framework introduced in [OW05] is something we could expect.

Another important model is given by Gatheral in [Gat10] who studies the concept of transient impact by introducing a decay factor. Gatheral deduces a relationship between the market impact function and its decay, assuming the next model for the evolution of market prices is

$$S_t := S_0 + \int_0^t h(\dot{x_s}) G(t-s) ds + \int_0^t \sigma dW_s \tag{6.14}$$

where $\dot{x}_s$ is the rate of trading at time $s < t$, $h(\dot{x}_s)$ corresponds to the temporary impact of the activity of the large investor at time s, the function $G(t-s)$ represents the decay factor and W_s is a Brownian motion. This functions $h(.)$ and $G(.)$ should be considered as representing the different market conditions in an average sense. We note that, for instance, in the framework of Almgren and Chriss given by [AC00], $G(t-s) = \delta(t-s)$ which means that the market impact decays instantaneously. In the case of [OW05], we have $G(t) = e^{-\rho t}$ with $\rho > 0$ and $h(.)$ is a linear function of the rate of trading.

6.2.5 A first extension of the second family of models

The work of [AFS10] extends the results in [OW05] by considering general shapes for the limit order book. In that paper, Alfonsi, Fruth and Schied assume that the process $(S_t^0)_{t\geq 0}$ is a rightcontinuous martingale on the filtered probability space $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P})$ such that $S_0^0 = S_0$ $\mathbb{P} - a.s.$. This unaffected price is not influenced by the large investor and so it is determined by the noise traders. The martingale assumption for the unaffected process $(S_t^0, t \geq 0)$ is also used in [BL98, AC00] and [OW05]. One of the reason that explains its use, is that as we consider short trading horizons, the drift effect can be neglected. As it is mentioned in [KSS11], a non-zero drift can create profitable round trips that would have to be differentiated from the ones defined by Huberman and Stanzl in [HS04].

Moreover, the aim of the model proposed in [AFS10] is to describe the dynamics of the price process $(S_t)_{t\geq 0}$ affected by the action of the large trader. This dynamic comes from the reaction of the limit order book and the dynamic of the unaffected price process. In order to describe the response of the limit order book, they introduce two processes: a volume impact process E_t that represents the impact of the large traders on the limit order book in terms of volume and a price impact process D_t that is the price impact created by the action of the large trader.

Thus, it is possible now to define the actual price process S as follows

$$S_t = S_t^0 + D_t. \tag{6.15}$$

Therefore, if at time t the trader places a market order of size ξ_t, the volume impact of the large investor changes from E_t to

$$E_{t+} := E_t + \xi_t \tag{6.16}$$

On the other hand, if the large trader is inactive, the process $(E_t, t \geq 0)$ decays exponentially as follows

$$dE_t = -\rho_t E_t dt \tag{6.17}$$

We consider that the rate at which E decreases is the deterministic and time dependent parameter $t \mapsto \rho_t$ called the resilience. The equations (6.16) and (6.17) define completely the dynamics of the volume impact process E and this framework is named the model with volume impact reversion.

The question is now, how to relate the volume impact and the price impact created by the order placed by the trader? In order to do that, they introduce a continuous distribution of bid and ask orders out of the unaffected price S_t^0. To represent this distribution, Alfonsi, Fruth and Schied consider a continuous function $f : \mathbb{R} \mapsto [0, \infty)$ verifying $f(x) > 0$ for a.e.x. This function is called the shape function of the limit order book. Thus, the number of shares offered between the prices $S_t^0 + x$ and $S_t^0 + x + dx$ is equal to $f(x)dx$.

Therefore, there is a relation that links the volume impact E_t with the price impact D_t, given by the following equation

$$\int_0^{D_t} f(x)dx = E_t. \tag{6.18}$$

Figure 6.8 shows the relation (6.18) between the volume impact E_t and the price impact D_t generated when a large trader executes an order on the market.

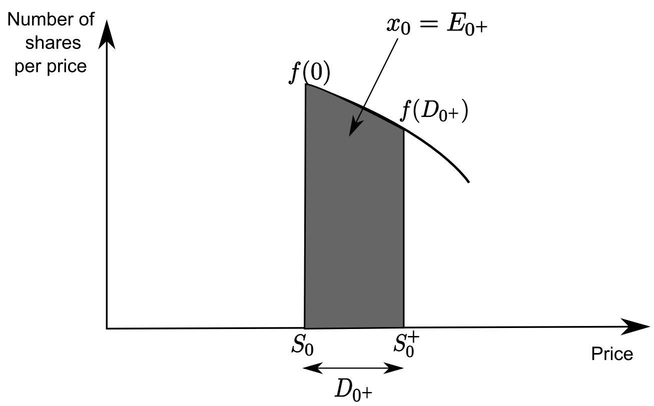


Fig. 6.7. The shape function f of the limit order book in the model [AFS10]. The trader executes an order of size E_{0+} and then moves the price from S_0 to S_{0+}.

Let us define the antiderivative of f,

$$F(y) := \int_0^y f(x)dx, \quad y \in \mathbb{R},$$

so we can observe that the connection between the volume impact process E and price impact process D can be stated as

$$E_t = F(D_t) \text{ and } D_t = F^{-1}(E_t) \tag{6.19}$$

We note that F is invertible because it is a strictly increasing function given the assumption $f(x) > 0$ for a.e. x.

We remark that if at time $t \geq 0$ the large investor places an order of size ξ_t, the actual price moves from S_t to

$$S_{t+} := S_t^0 + D_{t+} = S_t^0 + F^{-1}(E_t + \xi_t)$$

We can conclude that the price impact $D_{t+} - D_t$ is a non-linear function of the order size ξ_t except if the shape function of the limit order book f is constant between D_t and D_{t+}. This block-shaped limit order book is the one used, for instance, by Obizhaeva and Wang in [OW05] and corresponds to a linear price impact.

Alternatively, as we can consider that the resilience acts on the price impact when the large trader is not placing orders instead of the exponential reversion of the volume impact, by taking

$$dD_t = -\rho_t D_t dt \tag{6.20}$$

The equation (6.20) means that the price impact is recovering in the limit order book. So, if we replace (6.17) by (6.20) we obtain the so called price impact reversion model.

This two different models lead to different point of views for the dynamic of the limit order book and we remark that they are not necessarily equivalent. Thereby, instead of only studying the model impact with price reversion as in the most part of the literature, we also need to analyze the model with volume impact.

Instead of defining the optimal execution problem within a class of deterministic strategies as in [OW05], Alfonsi, Fruth and Schied in [AFS10] allow the strategy to belong to the larger class of adapted strategies. In [AFS10] the large trader could a priori have sell orders in a buy program, even though they show that this is not optimal in their model, alternatively to the model suggested in [OW05]. Moreover, analyzing several examples, Alfonsi, Fruth and Schied in [AFS10] assure the robustness of the optimal strategy.

The techniques used to solve the optimal execution problem in [AFS10] are based on the Lagrange multiplier method and not in dynamic programming techniques as in [OW05]. The idea behind is to reduce the model with the two sides of the limit order book to a model without bid-ask spread and to prove that these problems are equivalent.

The main results in the works [AFS10] and [AS10] are the existence and uniqueness of the optimal strategy and the fact that it is possible to exhibit analytic formulas for this optimal strategy.

6.2.6 Differences between the Gatheral model and the Alfonsi, Fruth and Schied model

One of the results obtained by [Gat10] is that the exponential decay is incompatible with the non-linear market impact, outcome that seems to be in contradiction with the conclusion in [AS10].

This difference can be explained by taking the discrete version of the model proposed by Gatheral in [Gat10]

$$S_t^G = S_t^0 + \sum_{\tau_n < t} h(\xi_n) G(t - \tau_n) \tag{6.21}$$

where the functions h and G are defined according to (6.14).

In the continuous-time case (6.14), considering an exponential decay $G(t) = e^{-\rho t}$, for ρ a positive constant, in [Gat10] Gatheral proved that if the impact function h is not linear there will not exist price manipulation strategies in the sense of Huberman and Stanzl. This conclusion can be also applied in the discrete version of the model presented in [Gat10] given by (6.21) using discrete approximations of the manipulation strategies obtained in the continuous case.

In the model proposed by Alfonsi, Fruth and Schied in [AFS10], if we consider a constant resilience ρ in the model with volume impact reversion, we obtain that the volume impact process is defined as follows

$$E_t = \sum_{\tau_n < t} \xi_n e^{-\rho(t-\tau_n)}, \tag{6.22}$$

and (6.22) implies that the actual price is given by

$$S_t = S_t^0 + F^{-1}\left(\sum_{\tau_n < t} \xi_n e^{-\rho(t-\tau_n)}\right). \tag{6.23}$$

Thus, the models proposed in [Gat10] and in [AFS10] are different which explains the difference in their conclusion because in (6.21) the nonlinear price impact function is applied to each trade whereas in (6.23) is a function of the total volume impact.

6.2.7 Price manipulation strategies

As an asset pricing model, a market impact model needs some requirements to assure its viability. Huberman and Stanzl in [HS04] introduce the concept of price manipulation, i.e. the existence of strategies where $\mathbb{x} = 0$ and that have strictly negative expected execution costs. These types of price manipulation strategies are called price manipulation strategies in the sense of Huberman and Stanzl. The presence of these price manipulations can be seen as a weak principle of no-quasi-arbitrage because the gain is in an average sense. Thus, by repeating indefinitely the price manipulation strategies, we can obtain by a law of large numbers, an almost sure profit, that is an arbitrage in the classical sense. So, Huberman and Stanzl affirm that the assumption of risk neutral dynamics for the unaffected price process is not sufficient for the viability of the market impact model because can exist price manipulation strategies for price impact models with martingale dynamics for prices. Besides, the problem that consists in the minimization of the costs incurred by the large trader could be not well defined due to the existence of the price manipulation strategies in the sense of Huberman and Stanzl. For example, by studying the relationship between the market impact function h and its decay G that exclude price manipulation strategies in the sense of Huberman and Stanzl [HS04], Gatheral obtains expressions determining the value of the parameters appearing in the price impact and decay functions.

Alfonsi and Schied in [AS10] extend the concept of price manipulation strategies by introducing the existence of strategies where the expected cost of a sell (buy) program can be decreased by intermediate buy (sell) trades, i.e. $\exists X$ admissible strategy such that

$$\mathcal{C}(X) < \inf \left\{ \mathcal{C}(\tilde{X}), \tilde{X} \text{ is admissible and nonincreasing or nondecreasing} \right\}.$$

Alfonsi and Schied called this notion the transaction-triggered price manipulation strategies. We remark that if there is no transaction-triggered price manipulation strategies then there is no price manipulation strategies in the sense of Huberman and Stanzl.

6.3 Motivation for our work

After overviewing all these models, we can note that they do not consider entirely the liquidity changes during the trading time $[0, T]$ where the large trader has to liquidate his position. Some of the models presented above consider a resilience parameter that depends on time but the distribution of orders in the limit order book does not depend on time and sometimes is even constant with respect to the price. Therefore, the aim of this part of the work is to extend the model proposed by [AS10] by introducing the time dependence of the liquidity and then by studying the consequences of this time component in markets while maintaining the tractability of the model introduced in [AFS10].

One work that considers that the liquidity is not constant over time is the paper of Fruth, Schöneborn and Urusov [FSU13]. Here, they analyze the optimal strategies and the existence of price manipulation strategies in a time-dependent limit order book model with deterministic depth and resilience for a non-zero and zero bid-ask spread case. In the case of a bid-ask spread different from zero, this spread increases when market orders are executed in the limit order book and so they prove the non existence of price manipulation strategies while they obtain that there are price manipulation strategies when the bid-ask spread is neglected.

Fruth, Schöneborn and Urusov in [FSU13] are interested in the minimization of costs problem which is set in a discrete-time and continuous-time framework and is solved using dynamic programming techniques. They find that under some assumptions on the parameters of the model, the orders have to be executed according to a ratio between the number of orders remaining to liquidate and the current price impact. We remark that the time-dependent limit order book used in [FSU13] is based on the model introduced in [OW05] which means that they take into account a block-shaped limit order book which implies considering a linear market impact.

In [BF12], Bank and Fruth also study the problem of letting the liquidity to change over time through the market depth and the resilience parameter. By considering also a block-shaped limit order book, they set out the minimization of costs as a convex problem in order to manipulate convex analysis techniques that allow to describe the optimal strategies using concave envelopes depending on the parameters of the model.

The work that we propose extends different points from the works [AFS10] and [AS10]. One new aspect is that the depth of the limit order book changes along with time. It is well known that there is more activity on markets at the opening and closing of markets but less activity at noon. This led to underlining the existence of an U-shaped pattern in the intraday trading volume, price volatility and average bid-ask spread, pattern that shows the relevance of time in the dynamic of the limit order book. These intraday patterns have been highlighted by Jain and Joh in [JJ88], Gerety and Mulher in [GM92] for volumes and prices in the U.S. markets, Hamao and Hasbrouck in [HH95] for the Tokyo Stock Exchange and Kleidon and Werner [KW96] for the London Stock Exchange. Hence, the idea of our model is to consider a time-dependent shape of the limit order book which is coherent with this time shift of the market conditions. For example, this shape could be calibrated to capture the deterministic changes in the liquidity during one day. See Chordia, Roll and Subrahmanyam [CRS01], Kempf and Mayston [KM08] and Lorenz and Osterrieder [LO09] for more details on deterministic liquidity patterns.

We remark that this calibration of the time-varying liquidity is consistent with the estimation of the other parameters of the model as the resilience because this resilience has to capture as well the changes in the recovery of the limit order book during a period of time.

Beyond solving the optimal execution problem in a more general context, our goal is to understand how the dynamics of the LOB may create or not price manipulations in this time-dependent framework. The study of this price manipulations strategies lead us to derive sufficient conditions to exclude them. These conditions are not only interesting from a theoretical point of view, as they also give a qualitative understanding on how price manipulations may occur when posting or cancelling limit orders. The behavior of market makers can be related to the parameters of the model which allow to explain these price manipulations strategies from the market maker point of view. It is important to note that price manipulations strategies do not exist in the model without time-dependence as it has been proven by [AFS10].

Another contribution of this work is that we solve the optimal execution problem in a continuous time setting while [AFS10] and [AS10] mainly focus on discrete time strategies. In particular, using a time-continuous framework is much more suitable for stating the conditions that exclude price manipulations and for understanding the influence of the shape of the limit order book on the price manipulations strategies.

Finally, we consider that new arrival orders can appear anywhere in the limit order book and not only between the bid-ask spread as it is considered in [AFS10] and [AS10]. In other words, the agents (noisy traders) can place limit orders in the depth of the limit order book and not only at the best price that is either the current best ask or the current best bid.

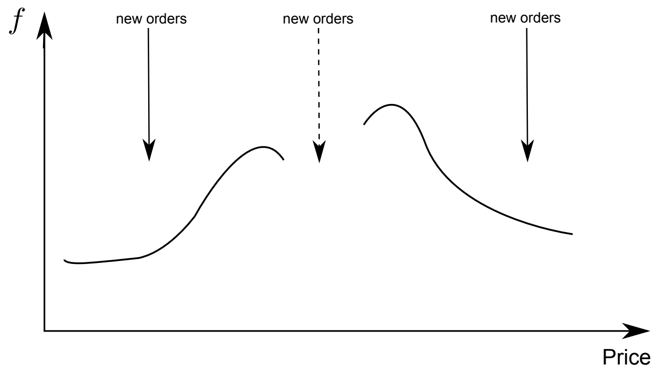


Fig. 6.8. New orders appear in the bid-ask spread in the model [AS10] as it is shown by the dashed line. In our framework, new orders can appear in the bid-ask spread but also in the depth of limit order book as can be seen with the continuous arrows.

Section 7 presents the model that extends [AFS10].

7

Optimal execution and price manipulations in time-varying limit order books

The paper that has been accepted in Applied Mathematical Finance is presented in this chapter.

Abstract: This paper focuses on an extension of the Limit Order Book (LOB) model with general shape introduced by Alfonsi, Fruth and Schied [AFS10]. Here, the additional feature allows a time-varying LOB depth. We solve the optimal execution problem in this framework for both discrete and continuous time strategies. This gives in particular sufficient conditions to exclude Price Manipulations in the sense of Huberman and Stanzl [HS04] or Transaction-Triggered Price Manipulations (see Alfonsi, Schied and Slynko [ASS11]). These conditions give interesting qualitative insights on how market makers may create or not price manipulations.

Introduction

It is a rather standard assumption in finance to consider an infinite liquidity. By infinite liquidity, we mean here that the asset price is given by a single value, and that one can buy or sell any quantity at this price without changing the asset price. This assumption is in particular made in the Black and Scholes model [BS73], and is often made as far as derivative pricing is concerned. When considering portfolio over a large time horizon, this approximation is relevant since one may split orders in small ones along the time and reduces one's own impact on the price. At most, the lack of liquidity can be seen as an additional transaction cost. This issue has been broadly investigated in the literature, see Cetin, Jarrow and Protter [ÇJP04] and references within.

If we consider instead brokers that have to trade huge volumes over a short time period (some hours or some days), we can no longer neglect the price impact of trading strategies. We have to focus on the market microstructure and model how prices are modified when buy and sell orders are executed. Generally speaking, the quotation of an asset is made through a Limit Order Book (LOB) that lists all the waiting buy and sell orders on this asset. The order prices have to be a multiple of the tick size, and orders at the same price are arranged in a First-In-First-Out stack. The bid (resp. ask) price is the price of the highest waiting buy (resp. lowest selling buy) order. Then, it is possible to buy or sell the asset in two different ways: one can either put a limit order and wait that this order matches another one or put a market order that consumes the cheapest limit orders in the book. In the first way, the transaction cost is known but the execution time is uncertain. In the second way, the execution is immediate (provided that the book contains enough orders). The price per share instead depends on the order size. For a buy (resp. sell) order, the first share will be traded at the ask (resp. bid) price while the last one will be traded some ticks upper (resp. lower) in order to fill the order size. The ask (resp. bid) price is then modified accordingly.

The typical issue on a short time scale is the optimal execution problem: on given a time horizon, how to buy or sell optimally a given amount of assets? As pointed in Gatheral [Gat10] and Alfonsi, Schied and Slynko [ASS11], this problem is closely related to the market viability and to the existence of price manipulations. Modelling the full LOB dynamics is not a trivial issue, especially if one wants to keep tractability to solve then the optimal execution problem. Instead, simpler models called market impact models have been proposed. These models only describe the dynamics of one asset price and model how the asset price is modified by a trading strategy. Thus, Bertsimas and Lo [BL98], Almgren and Chriss [AC00], Obizhaeva and Wang [OW05] have proposed different models where the price impact is proportional to the trading size, in which they solve the optimal execution problem. However, some empirical evidence on the markets show that the price impact of a trade is not proportional to its size, but is rather proportional to a power of its size (see for example Potters and Bouchaud [BP03], and references within). With this motivation in mind, Gatheral [Gat10] has suggested a nonlinear price impact model. In the same direction, Alfonsi, Fruth and Schied [AFS10] have derived a price impact model from a simple LOB modelling. Basically, the LOB is modelled by a shape function that describes the density of limit orders at a given price. This model has then been studied further by Alfonsi and Schied [AS10] and Predoiu, Shaikhet and Shreve [PSS11].

The present paper extends this model by letting the LOB shape function vary along the time. Beyond solving the optimal execution problem in a more general context, our goal is to understand how the dynamics of the LOB may create or not price manipulations. Indeed, a striking result in [AFS10, AS10] is that the optimal execution strategy is made with trades of same sign, which excludes any price manipulation. This result holds under rather general assumptions on the LOB shape function, when the LOB shape does not change along the time. Instead, we will see in this paper that a time-varying LOB may induce price manipulations and we will derive sufficient conditions to exclude them. These conditions are not only interesting from a theoretical point of view. They give a qualitative understanding on how price manipulations may occur when posting or cancelling limit orders. While preparing this work, Fruth, Schöneborn and Urusov [FSU13] have presented a paper where this issue is addressed for a block-shaped LOB, which amounts to a proportional price impact.

Here, we get back their result and extend it to general LOB shapes and thus nonlinear price impact. The other contribution of this paper is that we solve the optimal execution in a continuous time setting while [AFS10, AS10] mainly focus on discrete time strategies. This is in particular much more suitable to state the conditions that exclude price manipulations.

7.1 Market model and the optimal execution problem

7.1.1 The model description

The problem that we study in this paper is the classical optimal execution problem. To deal with this problem, we consider in this paper a framework which is a natural extension of the model proposed in Alfonsi, Fruth and Schied [AFS10] and developed by Alfonsi and Schied [AS10] and Predoiu, Shaikhet and Shreve [PSS11]. The additional feature that we introduce here is to allow a time varying depth of the order book. We consider a large trader who wants to liquidate a portfolio of $\mathtt{x}$ shares in a time period of $[0,T]$. In order to liquidate these $\mathtt{x}$ shares, the large trader uses only market orders, that is buy or sell orders that are immediately executed at the best available current price. Thus, our large trader cannot put limit orders. A long position $\mathtt{x} > 0$ will correspond to a sell program while a short position $\mathtt{x} < 0$ will stand for a a buy strategy. The optimal execution problem consists in finding the optimal trading strategy that minimizes the expected cost of the large trader.

We assume that the price process without the large trader would be given by a rightcontinuous martingale $(S_t^0, t \geq 0)$ on a given filtered probability space $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P})$. The actual price process $(S_t, t \geq 0)$ that takes into account the trades of the large trader is defined by:

$$S_t = S_t^0 + D_t,\ t \geq 0. \tag{7.1}$$

Thus, the process $(D_t, t \geq 0)$ describes the price impact of the large trader. We also introduce the process $(E_t, t \geq 0)$ that describes the volume impact of the large trader. If the large trader puts a market order of size ξ_t ($\xi_t > 0$ is a buy order and $\xi_t < 0$ a sell order), the volume impact process changes from E_t to:

$$E_{t+} := E_t + \xi_t. \tag{7.2}$$

When the large trader is not active, its volume impact E_t goes back to 0. We assume that it decays exponentially with a deterministic time-dependent rate $\rho_t > 0$ called resilience, so that we have:

$$dE_t = -\rho_t E_t dt. \tag{7.3}$$

We now have to specify how the processes D and E are related. To do so, we suppose a continuous distribution buy and sell limit orders around the unaffected price S_t^0: for $x \in \mathbb{R}$, we assume that the number of limit orders available between prices $S_t^0 + x$ and $S_t^0 + x + dx$ is given by $\lambda(t)f(x)dx$. These orders are sell orders if $x \geq D_t$ and buy orders otherwise. The functions $f : \mathbb{R} \mapsto (0, \infty)$ and $\lambda : [0, T] \mapsto (0, \infty)$ are assumed to be continuous, and represent respectively the LOB shape and the depth of the order book. We define the antiderivative of the function f, $F(y) := \int_0^y f(x)dx, \quad y \in \mathbb{R}$, and assume that

$$\lim_{x \to -\infty} F(x) = -\infty \text{ and } \lim_{x \to \infty} F(x) = \infty, \tag{7.4}$$

which means that the book contains an infinite number of limit buy and sell orders. Thus, we set the following relationship between the volume impact E_t and the price impact D_t:

$$\int_0^{D_t} \lambda(t) f(x) dx = E_t,$$

or equivalently,

$$E_t = \lambda(t) F(D_t) \text{ and } D_t = F^{-1}\left(\frac{E_t}{\lambda(t)}\right). \tag{7.5}$$

Within this framework, a large trade ξ_t changes D_t to $D_{t+} = F^{-1}\left(\frac{E_t + \xi_t}{\lambda(t)}\right)$ and has the cost

$$\int_{D_t}^{D_{t+}} (S_t^0 + x)\lambda(t) f(x) dx = \xi_t S_t^0 + \int_{D_t}^{D_{t+}} \lambda(t) x f(x) dx := \pi_t(\xi_t). \tag{7.6}$$

Throughout the paper, we assume that λ is $\mathcal{C}^2$ and set $\eta_t = \frac{\lambda'(t)}{\lambda(t)}$. Thus, we have

$$\lambda(t) = \lambda(0) \exp\left(\int_0^t \eta_u du\right),$$

and $t \mapsto \eta_t$ is $\mathcal{C}^1$. Similarly, we assume that $t \mapsto \rho_t$ is $\mathcal{C}^1$.

Now, let us observe that we have assumed that the volume impact decays exponentially when the large trader is inactive. Other choices are of course possible, and a natural one would be to assume that the price impact decays exponentially

$$dD_t = -\rho_t D_t dt, \tag{7.7}$$

which amounts to assume that $dE_t = \eta_t E_t dt - \rho_t \lambda(t) f(F^{-1}(E_t/\lambda(t))) F^{-1}(E_t/\lambda(t)) dt$ by (7.5).

Definition 7.1. *The dynamics of "model V" with volume impact reversion is the one given by* (7.1)*,* (7.2)*,* (7.3) *and* (7.5)*. The dynamics of "model P" with price impact reversion is the one given by* (7.1)*,* (7.2)*,* (7.7) *and* (7.5)*. In both models, we assume that the market is at equilibrium at time* 0*, i.e.* $E_0 = D_0 = 0$.

Remark 7.1. *Though being simplistic, this model describes through ρ_t and $\lambda(t)$ the two different ways that market makers have to put (or cancel) limit orders: it is either possible to pile orders at an* *existing price or to put orders at a better price than the existing ones. Thus, $\lambda(t)$ describes how market makers pile or cancel orders while ρ_t describes the rate at which new orders appear at a better price. Basically, one may think these functions one-day periodic, with relative high values at the opening and the closing of the market and low values around noon. The particular case $\lambda \equiv 1$ corresponds to the model introduced by Alfonsi, Fruth and Schied [AFS10] for which new orders can only appear at a better price.*

Let us make some comments of the model. Obviously, limit order books have in practice a much more sophisticated dynamics than the one presented above. Somehow, the shape function f should be seen as an average of all the order books observed during one day, while $\lambda(t)$ should represent a one day periodic function which is proportional to the volume of limit orders in the book. The resilience ρ_t should also be considered as a one-day periodic function that should be proportional to the volume of new limit orders that are posted at the bid or ask price. This give a way to estimate all the parameters. However, the scope of our model is not to model precisely the limit order book at a microstructure level. It aims at modelling the price impact at a mesoscopic time scale. The limit order book interpretation is mainly used to justify equation (7.6) that relates the transaction costs to the shape function f. Thus, it seems to us more appropriate to estimate the parameters using the market price impact observations. Empirically, it has been observed that the price impact is proportional to some power of the order size (see for instance Potters and Bouchaud [BP03], Almgren et al. [AHLT05], Tóth et al. [BDdL+11]). This indicates the choice of $f(x) = |x|^\gamma$ which reproduces precisely such an impact. The parameter γ could be estimated like in the cited papers. Once γ is chosen, $\lambda(t)$ could be estimated in the same vein from empirical data on market impact. Also, ρ_t can be estimated from the average time needed for the decay of the temporary price impact.

7.1.2 The optimal execution problem, and price manipulation strategies

We focus on the optimal liquidation of a portfolio with $\mathbb{x}$ shares by a large trader who can place market orders over a period of time $[0, T]$. Thus, $\mathbb{x} > 0$ (resp. $\mathbb{x} < 0$) corresponds to a selling (resp. buying) strategy.

We first consider discrete strategies and assume that at most $N + 1$ trades can occur. An admissible strategy will be then described by an increasing sequence $\tau_0 = 0 \leq \cdots \leq \tau_N = T$ of stopping times and random variables $\xi_0, \ldots, \xi_N$ (ξ_i stands for the trading size at time τ_i) such that

- $\mathbb{x} + \sum_{i=0}^{N} \xi_i = 0$, i.e. the trader liquidates indeed $\mathbb{x}$ shares,
- ξ_i is $\mathcal{F}_{\tau_i}$-measurable,
- $\exists M \in \mathbb{R}, \forall 0 \leq i \leq N, \xi_i \geq M$, a.s.

The expected cost of an admissible strategy $(\boldsymbol{\xi}, \mathcal{T})$ with $\boldsymbol{\xi} = (\xi_0, \ldots, \xi_N)$ and $\mathcal{T} = (\tau_0, \ldots, \tau_N)$ is given by

$$\mathcal{C}(\boldsymbol{\xi}, \mathcal{T}) = \mathbb{E}\left[\sum_{i=0}^{N} \pi_{\tau_i}(\xi_i)\right], \tag{7.8}$$

where $\pi_{\tau_i}(\xi_i)$ stands for the cost of the i-th trade, and is defined by (7.6) in models V or P. Throughout the paper, we will assume that the goal of the large trader is to minimize his expected cost among the admissible strategies.

Remark 7.2. *The choice of minimizing the expected cost enables us to characterize explicitly the optimal strategy (see Theorems 7.3, 7.4, 7.5 and 7.6). Let us mention here that other targets could be considered for the trader. Typically, one may wish to include in the optimization program some risk aversion. This has been considered for instance by Almgren and Chriss [AC00] or Schied [Sch13]. However, in our model we will see that the optimal strategy that minimizes the expected cost is deterministic. Thus, the risk only stems from the fluctuations of S^0 during the execution and can be in practice neglected when the deadline T is short enough, which gives a justification of our choice.*

We also consider continuous time trading strategy and make the same assumptions as Gatheral et al. [GSS12]. An admissible strategy $(X_t)_{t\geq 0}$ is a stochastic process such that

- $X_0 = \mathbb{x}$ and $X_{T+} = 0$,
- X is $(\mathcal{F}_t)$-adapted and leftcontinuous,
- the function $t \in [0, T+] \mapsto X_t$ has finite and a.s. bounded total variation.

The process X_t describes the number of shares that remains to liquidate at time t. Thus, the discrete time strategy above corresponds to $X_t = \mathbb{x} + \sum_{i=0}^{N} \xi_i \mathbf{1}_{\tau_i < t}$, and the three assumptions on $(\boldsymbol{\xi}, \mathcal{T})$ precisely give the ones on X. Let us observe that processes E and D are also leftcontinuous since we have in model V (resp. model P):

$$dE_t = dX_t - \rho_t E_t dt, \text{ (resp. } dE_t = dX_t + \eta_t E_t dt - \rho_t \lambda(t) f(F^{-1}(E_t/\lambda(t))) F^{-1}(E_t/\lambda(t)) dt). \tag{7.9}$$

We want now to write the cost associated to the strategy X. To do so, we introduce the following notations

$$x \in \mathbb{R}, \tilde{F}(x) = \int_0^x y f(y) dy, \; G(x) = \tilde{F}\left(F^{-1}(x)\right), \tag{7.10}$$

so that $\pi_t(dX_t) = S_t^0 dX_t + \lambda(t)[G\left(\frac{E_t + dX_t}{\lambda(t)}\right) - G\left(\frac{E_t}{\lambda(t)}\right)]$. Since $G' = F^{-1}$, the cost of an admissible strategy is given by:

$$\mathcal{C}(X) = \mathbb{E}\left[\int_0^T \left[S_t^0 + F^{-1}\left(\frac{E_t}{\lambda(t)}\right)\right] dX_t + \sum_{t \leq T} \lambda(t) \left[G\left(\frac{E_t + \Delta X_t}{\lambda(t)}\right) - G\left(\frac{E_t}{\lambda(t)}\right) - F^{-1}\left(\frac{E_t}{\lambda(t)}\right) \Delta X_t\right]\right], \tag{7.11}$$

which coincides with (7.8) for discrete strategies. Here, $\Delta X_t = X_{t+} - X_t$ denotes the jump of X at time t (jumps are countable), and dX_t stands for the signed measure on $[0, T]$ associated to $(X_t, 0 \leq$

$t \leq T+$) (a jump ΔX_T induces a Dirac mass in T). If we introduce the continuous part of X, $X_t^c := X_t - \sum_{0 \leq s < t} \Delta X_s$, we can rewrite the cost as follows:

$$\mathcal{C}(X) = \mathbb{E}\left[\int_0^T \left[S_t^0 + F^{-1}\left(\frac{E_t}{\lambda(t)}\right)\right] dX_t^c + \sum_{t \leq T} S_t^0 \Delta X_t + \lambda(t)\left[G\left(\frac{E_t + \Delta X_t}{\lambda(t)}\right) - G\left(\frac{E_t}{\lambda(t)}\right)\right]\right]. \quad (7.12)$$

The optimal execution problem is in fact closely related to questions around market viability and arbitrage. We recall the definition of price manipulation strategies introduced by Huberman and Stanzl [HS04].

Definition 7.2. *A round trip is an admissible strategy X for $\mathbb{x} = 0$. A Price Manipulation Strategy (PMS) in the sense of Huberman and Stanzl is a round trip whose expected cost is negative, i.e. $\mathcal{C}(X) < 0$.*

Heuristically, if a PMS exists, it could be repeated indefinitely and would lead to a classical arbitrage (i.e. an almost sure profit) by a law of large numbers. However, it has been pointed in Alfonsi et al. [ASS11] the absence of PMS does not ensure the market stability. In fact, in some PMS free models, the optimal strategy to sell $\mathbb{x}$ shares consists in buying and selling successively a much higher amount of shares. To correct this, they introduce the following definition.

Definition 7.3. *A model admits transaction-triggered price manipulations (TTPM) if the expected cost of a sell (buy) program can be decreased by intermediate buy (sell) trades, i.e.*

$$\exists X \textit{ admissible}, \ \mathcal{C}(X) < \inf\left\{\mathcal{C}(\tilde{X}), \tilde{X} \textit{ is admissible and nonincreasing or nondecreasing}\right\}.$$

It is rather natural choice to exclude TTPM: in presence of TTPM a large trader would increase the traded volume to minimize its cost, which produce noise and may yield to instability. Besides, the absence of TTPM implies the absence of PMS. The optimal strategy for buying $\varepsilon > 0$ shares is made only with intermediate buy trades and has thus a positive cost. Thus, by some cost continuity that usually holds (this is the case for models V and P), any round trip has a nonnegative cost.

Remark 7.3. *It is possible to define a two-sided limit order book model like in Alfonsi, Fruth and Schied [AFS10] or Alfonsi and Schied ([AS10], Section 2.6). In such a model, bid and ask prices evolve as follows. A buy (resp. sell) order of the large trader shifts the ask (resp. bid) price and leaves the bid (resp. ask) price unchanged. When the large trader is idle, the shifts on the ask and bid prices goes back exponentially to zero, like in models V or P. As in [AFS10, AS10], the two-sided model coincides with the model presented here when the large trader puts only buy orders or only sell orders. In particular, the optimal strategies are the same in both models in absence of TTPM.*

7.2 Main results

The first focus of this paper is to extend the results obtained in Alfonsi et al. [AFS10, AS10] and obtain the optimal execution strategies for LOB with a time varying depth λ. Doing so, our goal is also to better understand how this time varying depth may create manipulation strategies. In fact, it was shown in [AFS10] and [AS10] for $\lambda \equiv 1$ that under some general assumptions on the shape function f, there is an optimal liquidation strategy which is made only with sell (resp. buy) orders when $\mathbb{x} > 0$ (resp. $\mathbb{x} < 0$). Thus, there is no PMS nor TTPM when the LOB shape is constant. This is a striking result, and one may wonder how this is modified by changing slightly the assumptions. In Alfonsi, Schied and Slynko [ASS11] is studied the case of a block-shaped LOB, where the resilience is not exponential so that the market has some memory of the past trades. Conditions on the market resilience are given to exclude PMS and TTPM. Analogously, we want to obtain here conditions on λ and ρ that rules out such strategies. This is not only interesting from a theoretical point of view. This will give also some noticeable qualitative insights for market makers. In fact, for a market maker who places and cancels significant limit orders, these conditions will indicate if he may or not create manipulation strategies.

Before showing the results, it is worth to make further derivations on the expected cost. Let us start with discrete strategies. By using the martingale property on S^0 and the assumptions on $\boldsymbol{\xi}$ made in Section 7.1.2, we can show easily like in [AS10] that

$$\mathcal{C}(\boldsymbol{\xi}, \mathcal{T}) = -S_0^0 \mathbb{x} + \mathbb{E}\left[\sum_{i=0}^{N} \int_{D_{\tau_i}}^{D_{\tau_i+}} \lambda(\tau_i) x f(x) dx\right].$$

Then, it is easy to check that $\sum_{i=0}^{N} \int_{D_{\tau_i}}^{D_{\tau_i+}} \lambda(\tau_i) x f(x) dx$ is a deterministic function of $(\boldsymbol{\xi}, \mathcal{T})$ in both volume impact reversion and price impact reversion models. We respectively denote by $C^V(\boldsymbol{\xi}, \mathcal{T})$ and $C^P(\boldsymbol{\xi}, \mathcal{T})$ this function and get:

$$\mathcal{C}(\boldsymbol{\xi}, \mathcal{T}) = -S_0^0 \mathbb{x} + \mathbb{E}\left[C^M(\boldsymbol{\xi}, \mathcal{T})\right], \tag{7.13}$$

where $M \in \{V, P\}$ indicates the model chosen. Thus, if the function $(\boldsymbol{x}, \boldsymbol{t}) \mapsto C^M(\boldsymbol{x}, \boldsymbol{t})$ has a unique minimizer on $\{(\boldsymbol{x}, \boldsymbol{t}) \in \mathbb{R}^N \times \mathbb{R}^{N+1}, \sum_{i=1}^{N} x_i = -\mathbb{x}, 0 = t_0 \leq \cdots \leq t_N = T\}$, the optimal strategy is deterministic and given by this minimizer. When λ is constant, it is shown in [AS10] that under some assumptions on f depending on the model chosen, the optimal time grid $\boldsymbol{t}^\star$ is homogeneous with respect to ρ, i.e. $\int_{t_i^\star}^{t_{i+1}^\star} \rho_s ds = \frac{1}{N} \int_0^T \rho_s ds$. Instead, there is no such a simple characterization for general λ, even in the block-shaped case. Thus, we will focus on optimizing the trading strategy $\boldsymbol{\xi}$ on a fixed time grid $\boldsymbol{t}$:

$$\boldsymbol{t} = (t_0, \ldots, t_N), \text{ such that } 0 = t_0 < \cdots < t_N = T. \tag{7.14}$$

Last, we introduce the following notations that will be used throughout the paper:

$$a_i = e^{-\int_{t_{i-1}}^{t_i} \rho_u du}, \ \tilde{a}_i = \frac{a_i \lambda(t_{i-1})}{\lambda(t_i)} = e^{-\int_{t_{i-1}}^{t_i} (\rho_u + \eta_u) du}, \ \hat{a}_i = a_i \frac{\lambda(t_i)}{\lambda(t_{i-1})} = e^{-\int_{t_{i-1}}^{t_i} (\rho_u - \eta_u) du}, \quad 1 \le i \le N. \tag{7.15}$$

Similarly in the continuous case, we get by using the martingale assumption (see Lemma 2.3 in Gatheral, Schied and Slynko [GSS12]) that $\mathbb{E}[\int_0^T S_t^0 dX_t] = -\mathtt{x} S_0^0$. From (7.11) and (7.12), we get $\mathcal{C}(X) = -\mathtt{x} S_0^0 + \mathbb{E}[C^M(X)]$, where

$$C^M(X) = \int_0^T F^{-1}\left(\frac{E_t}{\lambda(t)}\right) dX_t^c + \sum_{t \le T} \lambda(t) \left[G\left(\frac{E_t + \Delta X_t}{\lambda(t)}\right) - G\left(\frac{E_t}{\lambda(t)}\right)\right].$$

Once again, C^M is a deterministic function of the strategy X in both models $M \in \{V, P\}$, and it is sufficient to focus on its minimization.

7.2.1 The block-shaped limit order book case ($f \equiv 1$)

In this section, we consider a shape function of the limit order book that has the form $\lambda(t)$. This time-dependent framework generalizes the block-shaped limit order book case studied by Obizhaeva and Wang [OW05] that consists in considering a uniform distribution of shares with respect to the price. We will get an explicit solution for the optimal execution problem, which extends the results given by Alfonsi, Fruth and Schied [AFS08].

Volume impact reversion model

When $f \equiv 1$, the deterministic cost function is simply given by

$$C^V(\boldsymbol{\xi}, \boldsymbol{t}) = \sum_{n=0}^N \lambda(t_n) \int_{D_{t_n}}^{D_{t_n+}} x f(x) dx = \sum_{i=0}^N \frac{\xi_i}{2} \left(\frac{\xi_i}{\lambda(t_i)} + 2 \frac{\sum_{j<i} e^{-\int_{t_j}^{t_i} \rho_s ds} \xi_j}{\lambda(t_i)} \right), \tag{7.16}$$

which is a quadratic form: $C^V(\boldsymbol{\xi}, \boldsymbol{t}) = \frac{1}{2} \boldsymbol{\xi}^T M^V \boldsymbol{\xi}$, with $M_{i,j}^V = \frac{\exp\left(-\left|\int_{t_i}^{t_j} \rho_s ds\right|\right)}{\lambda(t_i \vee t_j)}$, $0 \le i, j \le N$.

Theorem 7.1. *The quadratic form* (7.16) *is positive definite if and only if*

$$a_i \tilde{a}_i < 1, \forall i \in \{1, \dots, N\}. \tag{7.17}$$

In this case, the optimal execution problem to liquidate $\mathtt{x}$ *shares on the time-grid* (7.14) *admits a unique optimal strategy* $\boldsymbol{\xi}^\star$ *which is deterministic and explicitly given by:*

$$\begin{cases} \xi_0^\star &= -\frac{\mathtt{x}}{K_V} \lambda(t_0) \frac{1 - a_1}{1 - a_1 \tilde{a}_1} \\ \xi_i^\star &= -\frac{\mathtt{x}}{K_V} \lambda(t_i) \left[\frac{a_{i+1}}{1 - a_{i+1} \tilde{a}_{i+1}} (\tilde{a}_{i+1} - 1) + \frac{1 - \tilde{a}_i}{1 - a_i \tilde{a}_i} \right], \ 1 \le i \le N-1 \\ \xi_N^\star &= -\frac{\mathtt{x}}{K_V} \lambda(t_N) \frac{1 - \tilde{a}_N}{1 - a_N \tilde{a}_N}, \end{cases} \tag{7.18}$$

where

$$K_V = \frac{\lambda(t_0)(1-2a_1)+\lambda(t_1)}{1-a_1\tilde{a}_1} + \sum_{i=2}^{N} \lambda(t_i)\frac{(1-\tilde{a}_i)^2}{1-a_i\tilde{a}_i}.$$

Its cost is given by $C^V(\boldsymbol{\xi}^\star, \boldsymbol{t}) = \mathrm{x}^2/(2K_V)$.

This theorem provides an explicit optimal strategy for the large trader. It also gives explicit conditions that exclude or create PMS. First, let us assume that

$$\forall t \geq 0,\ 2\rho_t + \eta_t \geq 0. \tag{7.19}$$

Then, for any time grid (7.14), $a_i\tilde{a}_i \leq 1$ and the quadratic form (7.16) is positive semidefinite since it is a limit of positive definite quadratic forms. Thus, the model is PMS free. Conversely let us assume that $2\rho_{t_1} + \eta_{t_1} < 0$ for some $t_1 \geq 0$. Let us consider the following round trip on the time grid $\mathbf{t} = (0, t_1, t_2)$ with $t_2 > t_1$, where the large trader buys $x > 0$ at time t_1 and sells x at time t_2. The cost of such a strategy is given by

$$C^V((0,x,-x),\boldsymbol{t}) = \frac{x^2}{2\lambda(t_2)}\left(e^{\int_{t_1}^{t_2}\eta_u du} + 1 - 2e^{-\int_{t_1}^{t_2}\rho_u du}\right) \underset{t_2\to t_1}{=} \frac{x^2}{2\lambda(t_1)}\left((2\rho_{t_1}+\eta_{t_1})(t_2-t_1) + o(t_2-t_1)\right) \tag{7.20}$$

and is negative when t_2 is close enough to t_1.

Corollary 7.1. *In a block-shaped LOB, model V does not admit price manipulation in the sense of Huberman and Stanzl if and only if* (7.19) *holds.*

Let us now discuss this result from the point of view of market makers. A market maker that puts a significant amount of orders may have an influence on ρ_t and η_t and can increase (resp. decrease) them by respectively adding (resp. canceling) an order at a better price or at an existing limit order price. What comes out from (7.19) is that no PMS may arise if one adds limit orders, whatever the way of adding new orders. Instead, PMS can occur when canceling orders. A different conclusion will hold in the price reversion model.

An analogous result to Corollary 7.1 is stated in a recent paper by Fruth, Schöneborn and Urusov [FSU13] that has been posted while we were preparing this work. To be precise, results in [FSU13] are given for model P with a block-shaped LOB, and the optimal execution strategy is obtained in a continuous time setting. As we will see in the next paragraph, models V and P are mathematically equivalent when the LOB shape is constant, even though they are different from a financial point of view. By taking a regular time-grid $t_i = \frac{iT}{N}, i = 0\ldots, N$, and letting $N \to +\infty$, we get back the optimal strategy in continuous time (that we still denote by $\xi^\star$, by a slight abuse of notations):

$$\begin{cases} \xi_0^\star \underset{N\to+\infty}{\longrightarrow} \xi_0^\star := -\frac{\mathbb{x}}{\lambda(T)+\int_0^T \frac{\rho_s^2\lambda(s)}{\eta_s+2\rho_s}ds}\frac{\lambda(0)\rho_0}{2\rho_0+\eta_0} \\ \frac{\xi_{i_N}^\star}{T/N} \underset{N\to+\infty}{\longrightarrow} \xi_t^\star := -\frac{\mathbb{x}}{\lambda(T)+\int_0^T \frac{\rho_s^2\lambda(s)}{\eta_s+2\rho_s}ds}\lambda(t)\left[\left(\frac{\rho_t}{2\rho_t+\eta_t}\right)' + \rho_t\left(\frac{\rho_t+\eta_t}{2\rho_t+\eta_t}\right)\right], \text{ for } i_N \text{ such that } \frac{t_{i_N}}{N} \underset{\Delta t\to 0}{\longrightarrow} t \\ \xi_N^\star \underset{N\to+\infty}{\longrightarrow} \xi_T^\star := -\frac{\mathbb{x}}{\lambda(T)+\int_0^T \frac{\rho_s^2\lambda(s)}{\eta_s+2\rho_s}ds}\frac{\lambda(T)(\eta_T+\rho_T)}{\eta_T+2\rho_T}. \end{cases} \tag{7.21}$$

The strategy $dX_t^\star = \xi_0^\star\delta_0(dt) + \xi_t^\star dt + \xi_T^\star\delta_T(dt)$ with initial trade $\xi_0^\star$, continuous trading $\xi_t^\star$ on $[t, t+dt]$ for $t \in (0,T)$ and last trade $\xi_T^\star$ is indeed shown to be optimal in Fruth, Schöneborn and Urusov [FSU13] among the continuous time strategies with bounded variation. We will show here again this result for more general LOB shape. The optimal strategy has the following cost:

$$\frac{\mathbb{x}^2}{2\left[\lambda(T) + \int_0^T \frac{\rho_s^2\lambda(s)}{2\rho_s+\eta_s}ds\right]}.$$

Besides, this provides a necessary and sufficient condition to exclude transaction-triggered price manipulation.

Corollary 7.2. *In a block-shaped LOB, model V does not admit transaction-triggered price manipulation if and only if*

$$\forall t \geq 0,\ \eta_t + \rho_t \geq 0,\ \textit{and}\ \left(\frac{\rho_t}{2\rho_t+\eta_t}\right)' + \rho_t\left(\frac{\rho_t+\eta_t}{2\rho_t+\eta_t}\right) \geq 0. \tag{7.22}$$

The first condition comes from the last trade and implies (7.19) since $\rho_t \geq 0$. It can be interpreted similarly as condition (7.19) from market makers' point of view. The second condition in (7.22) comes from the intermediate trades and brings on the derivatives of ρ and η. It is harder to get an intuitive idea of its meaning from a market maker's point of view. Last, let us mention that we can show that the optimal strategy on the discrete time-grid (7.14) is made with nonnegative trades if one has (7.17) and

$$\frac{1-\tilde{a}_i}{1-a_i\tilde{a}_i} \geq a_{i+1}\frac{1-\tilde{a}_{i+1}}{1-a_{i+1}\tilde{a}_{i+1}},\ \ \forall i \in \{1,\dots,N-1\}\ \text{and}\ \tilde{a}_N \leq 1. \tag{7.23}$$

Condition (7.22) can be seen as the continuous time limit of condition (7.23).

Let us give now an illustration of the optimal strategy with a time-varying depth. We consider the case of a time-varying depth

$$\lambda(t) = \lambda_0 + \cos(2\pi t),\ \text{with}\ \lambda_0 > 1,$$

which corresponds to a one-day periodic function with high values at the beginning and at the end of the day. We can show that $\eta_t \geq -\frac{2\pi}{\sqrt{\lambda_0^2-1}}$ and with a constant resilience ρ, there is no PMS as soon as $2\rho - \frac{2\pi}{\sqrt{\lambda_0^2-1}} \geq 0$. Figure 7.1 shows the optimal execution strategy (7.18) with a value λ_0 that excludes PMS but allows TTPM. The optimal strategy to buy 50 shares consists in buying almost 95 shares and selling 45 shares, which roughly trebles the traded volume.

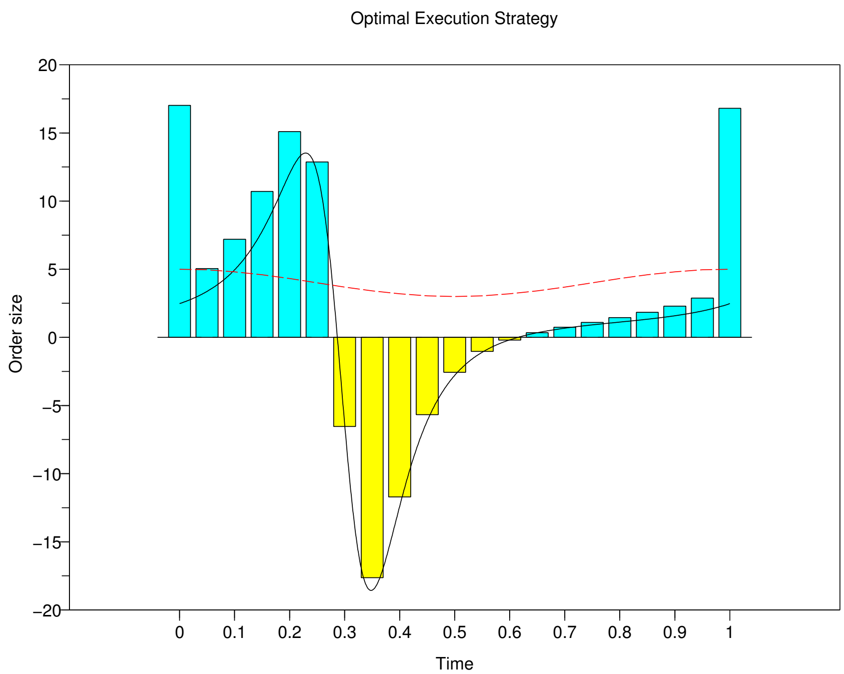


Fig. 7.1. Optimal execution strategy to buy 50 shares on a regular time grid, with $N = 20$, $\rho = 1$, $\lambda(t) = 4 + \cos(2\pi t)$ (plotted in dashed line). In solid line is plotted the function $t \mapsto \left(\frac{\rho_t}{2\rho_t+\eta_t}\right)' + \rho_t \left(\frac{\rho_t+\eta_t}{2\rho_t+\eta_t}\right)$.

Price impact reversion model

When $f \equiv 1$, the deterministic cost function $\sum_{i=0}^{N} \int_{D_{t_i}}^{D_{t_i+}} \lambda(t_i) x f(x) dx$ is given by

$$C^P(\boldsymbol{\xi}, \boldsymbol{t}) = \sum_{n=0}^{N} \lambda(t_n) \int_{D_{t_n}}^{D_{t_n+}} x f(x) dx = \sum_{i=0}^{N} \frac{\xi_i}{2} \left(\frac{\xi_i}{\lambda(t_i)} + 2 \sum_{j<i} e^{-\int_{t_j}^{t_i} \rho_s ds} \frac{\xi_j}{\lambda(t_j)} \right). \tag{7.24}$$

This is a quadratic form: $C^P(\boldsymbol{\xi}, \boldsymbol{t}) = \frac{1}{2} \boldsymbol{\xi}^T M^P \boldsymbol{\xi}$, with $M^P_{i,j} = \frac{\exp\left(-\left|\int_{t_j}^{t_i} \rho_s ds\right|\right)}{\lambda(t_i \wedge t_j)}$ for $0 \le i,j \le N$. When $f \equiv 1$, we get from (7.9) that model P is equivalent to model V with a resilience $\tilde{\rho}_t = \rho_t - \eta_t$. Another way to see that both models are mathematically the same in the block-shape case is to reverse the time and consider:

$$\forall t \in [0, T],\ \hat{\rho}_t = \rho_{T-t},\ \hat{\lambda}(t) = \lambda(T-t) \text{ and } \hat{t}_{N-i} = T - t_i, \text{ for } 0 \le i \le N.$$

Then, we have

$$M^P_{i,j} = \frac{e^{-|\int_{t_j}^{t_i} \rho_s ds|}}{\lambda(t_i \wedge t_j)} = \frac{e^{-|\int_{\hat{t}_{N-i}}^{\hat{t}_{N-j}} \hat{\rho}_s ds|}}{\hat{\lambda}(\hat{t}_{N-i} \vee \hat{t}_{N-j})}, \tag{7.25}$$

and the optimal execution problem in Model P with resilience ρ, LOB depth $\lambda(t)$ and time-grid $\boldsymbol{t}$ is the same as the optimal execution problem in Model V with resilience $\hat{\rho}$, LOB depth $\hat{\lambda}(t)$ and time-grid $\hat{\boldsymbol{t}}$. We immediately get the following results.

Theorem 7.2. *The quadratic form* (7.24) *is positive definite if and only if*

$$a_i\hat{a}_i < 1, \forall i \in \{1,\dots,N\} \tag{7.26}$$

In this case, the optimal execution problem to liquidate $\mathbb{x}$ *shares on the time-grid* (7.14) *admits a unique optimal strategy* $\boldsymbol{\xi}^\star$ *which is deterministic and explicitly given by:*

$$\begin{cases} \xi_0^\star = -\frac{\mathbb{x}}{K_P}\lambda(t_0)\frac{1-\hat{a}_1}{1-a_1\hat{a}_1}. \\ \xi_i^\star = -\frac{\mathbb{x}}{K_P}\lambda(t_i)\left[\frac{a_i}{1-a_i\hat{a}_i}(\hat{a}_i-1)+\frac{1-\hat{a}_{i+1}}{1-a_{i+1}\hat{a}_{i+1}}\right],\ 1\le i\le N-1 \\ \xi_N^\star = -\frac{\mathbb{x}}{K_P}\lambda(t_N)\frac{1-a_N}{1-a_N\hat{a}_N} \end{cases} \tag{7.27}$$

where

$$K_P = \frac{\lambda(t_N)(1-2a_N)+\lambda(t_{N-1})}{1-a_N\hat{a}_N} + \sum_{i=0}^{N-2}\lambda(t_i)\frac{(1-\hat{a}_{i+1})^2}{1-a_{i+1}\hat{a}_{i+1}}.$$

Its cost is given by $C^P(\boldsymbol{\xi}^\star,\boldsymbol{t}) = \mathbb{x}^2/(2K_P)$.

By taking a regular time-grid $t_i = \frac{iT}{N}, i = 0\dots,N$, and letting $N \to +\infty$, we get the optimal strategy in continuous time:

$$\begin{cases} \xi_0^\star \underset{N\to\infty}{\longrightarrow} \xi_0^\star := -\frac{\mathbb{x}}{\lambda(0)+\int_0^T \frac{\rho_s^2\lambda(s)}{2\rho_s-\eta_s}ds}\lambda(0)\frac{\rho_0-\eta_0}{2\rho_0-\eta_0} \\ \frac{\xi_{i_N}^\star}{T/N} \underset{N\to\infty}{\longrightarrow} \xi_t^\star := -\frac{\mathbb{x}}{\lambda(0)+\int_0^T \frac{\rho_s^2\lambda(s)}{2\rho_s-\eta_s}ds}\lambda(t)\left[\left(\frac{\rho_t-\eta_t}{2\rho_t-\eta_t}\right)'+\rho_t\left(\frac{\rho_t-\eta_t}{2\rho_t-\eta_t}\right)\right], \text{ for } i_N \text{ such that } \frac{Ti_N}{N} \underset{\Delta t\to 0}{\longrightarrow} t \\ \xi_N^\star \underset{N\to\infty}{\longrightarrow} \xi_T^\star := -\frac{\mathbb{x}}{\lambda(0)+\int_0^T \frac{\rho_s^2\lambda(s)}{2\rho_s-\eta_s}ds}\lambda(T)\frac{\rho_T}{2\rho_T-\eta_T}. \end{cases} \tag{7.28}$$

The strategy with initial trade $\xi_0^\star$, continuous trading $\xi_t^\star$ on $[t,t+dt]$ for $t\in(0,T)$ and last trade $\xi_T^\star$ is shown to be optimal in Fruth, Schöneborn and Urusov [FSU13] among the continuous time strategies with bounded variation, and has the following cost:

$$\frac{\mathbb{x}^2}{2\left[\lambda(0)+\int_0^T \frac{\rho_s^2\lambda(s)}{2\rho_s-\eta_s}ds\right]}.$$

Corollary 7.3. *In a block-shaped LOB, model* P *does not admit price manipulation in the sense of Huberman and Stanzl if and only if*

$$\forall t\ge 0,\ 2\rho_t-\eta_t\ge 0. \tag{7.29}$$

It does not admit transaction-triggered price manipulation if and only if

$$\forall t\ge 0,\ \rho_t-\eta_t\ge 0,\ \textit{and}\ \left(\frac{\rho_t-\eta_t}{2\rho_t-\eta_t}\right)'+\rho_t\left(\frac{\rho_t-\eta_t}{2\rho_t-\eta_t}\right)\ge 0. \tag{7.30}$$

The first condition in (7.30) comes from the initial trade while the second comes from intermediate trades. From market makers' point of view, (7.29) and the first condition in (7.30) give different conclusions from model V. A significant market maker will not create manipulation strategy if he puts orders at a better price (which increases ρ) or cancels orders at existing prices (which decreases η). Instead, he may create manipulation strategies if he piles orders at existing prices, or if he cancels orders that are among the best offers. The second condition of (7.30) brings on the dynamics of η and ρ and it is more difficult to give its heuristic meaning in terms of trading. Last, let us mention that the optimal strategy in discrete time given by Theorem 7.2 is made only with trades of same sign if, and only if, one has (7.26) and

$$\frac{1-\hat{a}_{i+1}}{1-a_{i+1}\hat{a}_{i+1}} \geq a_i \frac{1-\hat{a}_i}{1-a_i\hat{a}_i}, \quad \forall i \in \{1,\dots,N-1\} \text{ and } \hat{a}_1 < 1. \tag{7.31}$$

7.2.2 Results for general limit order book shape

We extend in this section the results obtained on the optimal execution for block-shaped LOB to more general shapes. In particular, the necessary and sufficient conditions that we have obtained to exclude TTPM (namely (7.22) for model V and (7.30) for model P) are still sufficient conditions to exclude TTPM for a wider class of shape functions. From a mathematical point of view, we proceed as follows. We first characterize the optimal strategy on a discrete time grid, by using Lagrange multipliers. Then, one can guess the optimal continuous time strategy, and we prove its optimality by a verification argument.

Volume impact reversion model

We first introduce the following assumption that will be useful to study the optimal discrete strategy.

Assumption 7.1. *1. The shape function f satisfies the following condition:*

$$f \textit{ is nondecreasing on } \mathbb{R}_- \textit{ and nonincreasing on } \mathbb{R}_+$$

2. $\forall t \geq 0, \rho_t + \eta_t \geq 0$.

We remark that when the LOB shape does not evolve in time ($\eta_t = 0$), the second condition is satisfied and we get back the assumption made in Alfonsi, Fruth and Schied [AFS10]. We define

$$x \in \mathbb{R},\ h_{V,i}(x) = \frac{F^{-1}(x) - a_i F^{-1}(\tilde{a}_i x)}{1-a_i}, 1 \leq i \leq N. \tag{7.32}$$

Theorem 7.3. *Under Assumption 7.1, the cost function $C^V(\boldsymbol{\xi}, \boldsymbol{t})$ is nonnegative, and there is a unique optimal execution strategy $\boldsymbol{\xi}^\star$ that minimizes C^V over $\{\boldsymbol{\xi} \in \mathbb{R}^{N+1}, \sum_{i=0}^N \xi_i = -\mathbb{x}\}$. This strategy is given as follows. The following equation*

$$\sum_{i=1}^N \lambda(t_{i-1})(1-a_i)h_{V,i}^{-1}(\nu) + \lambda(t_N)F(\nu) = -\mathbb{x}$$

has a unique solution $\nu \in \mathbb{R}$, and

$$\begin{aligned}
\xi_0^\star &= \lambda(t_0)h_{V,1}^{-1}(\nu)\,, \\
\xi_i^\star &= \lambda(t_i)(h_{V,i+1}^{-1}(\nu) - \tilde{a}_i h_{V,i}^{-1}(\nu)),\ \ 1 \le i \le N-1, \\
\xi_N^\star &= \lambda(t_N)F(\nu) - \lambda(t_{N-1})a_N h_{V,N}^{-1}(\nu)\,.
\end{aligned}$$

The first and the last trade have the same sign as $-\mathbb{x}$. Besides, if the following condition holds

$$\frac{1}{\tilde{a}_i}\frac{1-\tilde{a}_i}{1-a_i} \ge \frac{1-\tilde{a}_{i+1}}{1-a_{i+1}}, \tag{7.33}$$

the intermediate trades $\xi_i^\star$, $1 \le i \le N-1$, have also the same sign as $-\mathbb{x}$.

This theorem extends the results of [AFS10], where λ is assumed to be constant. In that case, (7.33) is satisfied and all the trades have the same sign. Condition (7.33) is interesting since it does not depend on the shape function, but it is more restrictive than the condition (7.23) for the block-shape case (see Lemma 7.4 for (7.33) $\implies$ (7.23)). In fact, the continuous time formulation is more convenient to analyze the sign of the trades. Under Assumption 7.1, we will show that no transaction-triggered price manipulation can occur with the same condition (7.22) as for the block-shape case.

When stating the optimal continuous-time strategy, we slightly relax Assumption 7.1. This is basically due to the argument of the proof that relies on a verification argument. Instead, our proof in the discrete case relies on Lagrange multipliers which requires to show first that the cost function has a minimum, and we use $\rho_t + \eta_t \ge 0$ for that. We introduce the following function

$$h_{V,t}(x) = F^{-1}(x) + \frac{\eta_t + \rho_t}{\rho_t}\frac{x}{f(F^{-1}(x))}. \tag{7.34}$$

We will show that no PMS exists and that there is a unique optimal strategy if these functions for $t \in [0,T]$ are bijective on $\mathbb{R}$ with a positive derivative. If Assumption 7.1 holds, this condition is automatically satisfied.

Theorem 7.4. *Let $f \in \mathcal{C}^1(\mathbb{R})$. We assume that for $t \in [0,T]$, $h_{V,t}$ is bijective on $\mathbb{R}$, such that $h'_{V,t} > 0$. Then, the cost function $C^V(X)$ is nonnegative, and there is a unique optimal admissible strategy $X^\star$ that minimizes C^V. This strategy is given as follows. The equation*

$$\int_0^T \lambda(t)\rho_t h_{V,t}^{-1}(\nu)dt + \lambda(T)F(\nu) = -\mathbb{x} \tag{7.35}$$

has a unique solution $\nu \in \mathbb{R}$ and we set $\zeta_t = h_{V,t}^{-1}(\nu)$. The strategy $dX_t^\star = \xi_0^\star\delta_0(dt) + \xi_t^\star dt + \xi_T^\star\delta_T(dt)$ with

$$\begin{aligned}
\xi_0^\star &= \lambda(0)\zeta_0, \\
\xi_t^\star &= \lambda(t)\left[\frac{d\zeta_t}{dt} + (\rho_t + \eta_t)\zeta_t\right], \\
\xi_T^\star &= \lambda(T)(F(\nu) - \zeta_T),
\end{aligned}$$

is optimal. The initial trade $\xi_0^\star$ has the same sign as $-\mathbb{x}$.

Thus, a sufficient condition to exclude price manipulation strategies is to assume that $h_{V,t}$ is bijective with $h'_{V,t} > 0$. We have a partial reciprocal result: there are PMS as soon as $h'_{V,t_1}(0) < 0$ (or equivalently, $2\rho_{t_1} + \eta_{t_1} < 0$) for some $t_1 \geq 0$. Indeed, in this case we consider the following round trip on the time grid $\mathbf{t} = (0, t_1, t_2)$ with $t_2 > t_1$, where the large trader buys $x > 0$ at time t_1 and sells x at time t_2. The cost of such a strategy is given by

$$\begin{aligned}
C^V((0,x,-x),\boldsymbol{t}) &= \lambda(t_1)G\left(\frac{x}{\lambda(t_1)}\right) + \lambda(t_2)\left(G\left(\frac{x(e^{-\int_{t_1}^{t_2}\rho_s ds} - 1)}{\lambda(t_2)}\right) - G\left(\frac{xe^{-\int_{t_1}^{t_2}\rho_s ds}}{\lambda(t_2)}\right)\right) \\
&= \lambda(t_1)\left(-\eta_{t_1}G\left(\frac{x}{\lambda(t_1)}\right) + (\rho_{t_1} + \eta_{t_1})\frac{x}{\lambda(t_1)}F^{-1}\left(\frac{x}{\lambda(t_1)}\right)\right)(t_2 - t_1) + o(t_2 - t_1).
\end{aligned}$$

The derivative of $x \mapsto -\eta_{t_1}G(x) + (\rho_{t_1} + \eta_{t_1})xF^{-1}(x)$ is $\rho_{t_1}h_{V,t_1}(x)$, which has the opposite sign of x near 0 since $h_{V,t_1}(0) = 0$ and $h'_{V,t_1}(0) < 0$ by assumption. Thus, we have $C^V((0,x,-x),\mathbf{t}) < 0$ for x and $t_2 - t_1$ small enough.

Now, let us focus on the sign of the trades given by the optimal strategy. Without further hypothesis, the condition $\xi_t^\star \geq 0$ typically involves the shape function f. However, under Assumption 7.1, we can show that TTPM are excluded under the same assumption as for the block-shape case.

Corollary 7.4. *Let $f \in \mathcal{C}^1$. Under Assumption 7.1, the function $h_{V,t}$ is $\mathcal{C}^1(\mathbb{R})$, bijective on $\mathbb{R}$, and such that $h'_{V,t} > 0$. Thus, the result of Theorem 7.4 holds and the last trade $\xi_T^\star$ has the same sign as $-\mathbb{x}$.*

Besides, if (7.22) *also holds, $\xi_t^\star$ has the same sign as $-\mathbb{x}$ for any $0 < t < T$, which excludes TTPM.*

Let us now focus on the example of a power-law shape: we assume that

$$f(x) = |x|^\gamma,\ \gamma > -1.$$

Strictly speaking, this does not fulfill the assumptions made before on f since it is not smooth in 0. However, the result of Theorem 7.4 still holds. In this case, $F(x) = \mathrm{sgn}(x)\frac{|x|^{\gamma+1}}{\gamma+1}$ is well-defined and satisfies (7.4). We have $F^{-1}(x) = \mathrm{sgn}(x)(\gamma+1)^{\frac{1}{\gamma+1}}|x|^{\frac{1}{\gamma+1}}$ and $h_{V,t}(x) = \mathrm{sgn}(x)(\gamma+1)^{\frac{1}{\gamma+1}}|x|^{\frac{1}{\gamma+1}}\left(\frac{\rho_t(2+\gamma)+\eta_t}{\rho_t(1+\gamma)}\right)$. Thus, $h_{V,t}$ is bijective and increasing if, and only if:

$$\rho_t(2+\gamma) + \eta_t > 0.$$

In this case, we have

$$h_{V,t}^{-1}(x) = \frac{1}{\gamma+1}K_t(\gamma)\mathrm{sgn}(x)|x|^{\gamma+1} \text{ with } K_t(\gamma) = \left(\frac{\rho_t(1+\gamma)}{\rho_t(2+\gamma)+\eta_t}\right)^{1+\gamma}.$$

In this case, we have by Theorem 7.4 that

$$\begin{cases} \xi_0^\star = \frac{-\mathrm{x}}{\int_0^T \lambda(t)\rho_t K_t(\gamma)dt + \lambda(T)}\lambda(0)K_0(\gamma), \\ \xi_t^\star = \frac{-\mathrm{x}}{\int_0^T \lambda(t)\rho_t K_t(\gamma)dt + \lambda(T)}\lambda(t)\left[\frac{dK_t(\gamma)}{dt} + (\rho_t+\eta_t)K_t(\gamma)\right] \\ \xi_T^\star = \frac{-\mathrm{x}}{\int_0^T \lambda(t)\rho_t K_t(\gamma)dt + \lambda(T)}\lambda(T)(1-K_T(\gamma)) \end{cases} \tag{7.36}$$

is the unique optimal strategy. For $\gamma = 0$, we get back (7.21). If we only assume that $\rho_t(2+\gamma)+\eta_t \geq 0$, we still have $C^V(X) \geq 0$ for any admissible strategy X. The cost $C^V(X)$ is indeed continuous with respect to the resilience, and is the limit of the cost associated to resilience $\rho_t + \varepsilon$, $\varepsilon \downarrow 0$. On the contrary, if $\rho_t(2+\gamma)+\eta_t < 0$, we have $h'_{V,t}(0) < 0$ and there is a PMS as explained above.

Corollary 7.5. *When $f(x) = |x|^\gamma$, model V does not admit PMS if, and only if*

$$\forall t \geq 0,\ \rho_t(2+\gamma) + \eta_t \geq 0.$$

It does not admit transaction-triggered price manipulation if and only if

$$\forall t \geq 0,\ \rho_t + \eta_t \geq 0,\ \textit{and}\ \left(\frac{\rho_t(1+\gamma)}{\rho_t(2+\gamma)+\eta_t}\right)' + \rho_t\left(\frac{\rho_t+\eta_t}{\rho_t(2+\gamma)+\eta_t}\right) \geq 0.$$

These conditions come respectively from the nonnegativity of the last and intermediate trades. For given functions ρ_t and η_t, the no PMS condition will be satisfied for $t \in [0,T]$ when γ is large enough. This can be explained heuristically. When γ increases, limit orders become rare close to S_t^0 and dense away from S_t^0, which creates some bid-ask spread. One has then to pay to get liquidity, and round trips have a positive cost. Instead, when γ is close to -1 it is rather cheap to consume limit orders, which may facilitate PMS. In Figure 7.2, we have plotted the optimal strategy for $\gamma = -0.3$ and $\gamma = 1$ with the same parameters as in Figure 7.1 for the Block shape case. We can check that the no PMS condition is satisfied in both cases.

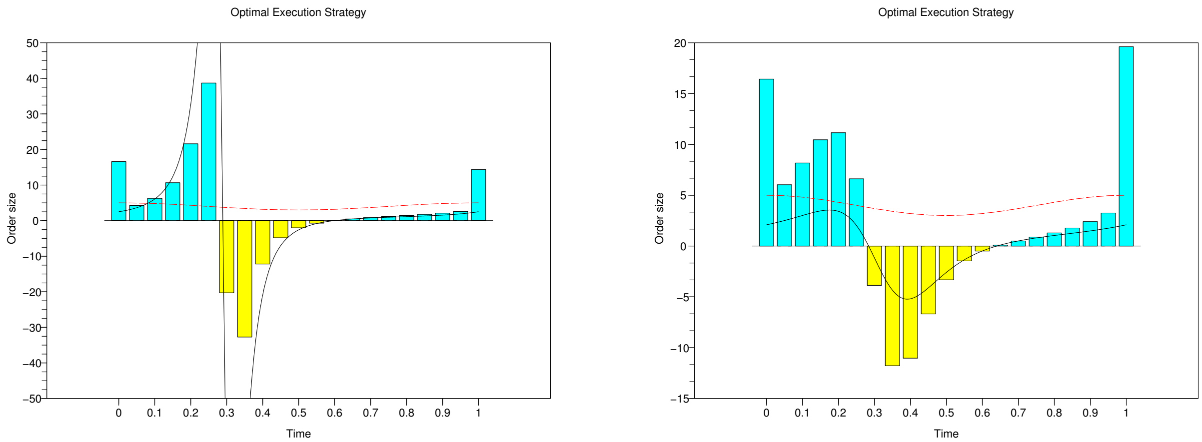


Fig. 7.2. Optimal execution strategy to buy 50 shares on a regular time grid, with $N = 20$, $\rho = 1$, $\lambda(t) = 4 + \cos(2\pi t)$ (plotted in dashed line) and $\gamma = -0.3$ (left) or $\gamma = 1$ (right). In solid line is plotted the function $t \mapsto \left(\frac{\rho_t(1+\gamma)}{\rho_t(2+\gamma)+\eta_t}\right)' + \rho_t\left(\frac{\rho_t+\eta_t}{\rho_t(2+\gamma)+\eta_t}\right)$ (this function is well-defined but out of the graph for $\gamma = -0.3$).

Price impact reversion model

The results that we present for model P are similar to the one obtained for model V. We first solve the optimal execution problem in discrete time. From its explicit solution, we then calculate its continuous time limit and check by a verification argument that it is indeed optimal. Doing so, we get sufficient conditions to exclude PMS and TTPM. In particular, condition (7.30) that excludes PMS and TTPM for block-shape LOB also excludes PMS and TTPM for a general LOB shape satisfying Assumption 7.2 below.

To study the optimal discrete strategy, we will work under the following assumption.

Assumption 7.2. *1. The shape function f is $\mathcal{C}^1$ and satisfies the following condition:*

$$f \text{ is nonincreasing on } \mathbb{R}_- \text{ and nondecreasing on } \mathbb{R}_+$$

2. $\forall t \geq 0, \rho_t - \eta_t > 0$.

3. $x \mapsto x\frac{f'(x)}{f(x)}$ *is nondecreasing on* $\mathbb{R}_-$*, nonincreasing on* $\mathbb{R}_+$.

The monotonicity assumption made here is the opposite to the one made in Assumption 7.1 for model V. This choice is different from the one made in Alfonsi et al. [AFS10, AS10]. It is in fact more tractable from a mathematical point of view, especially here with a time-varying LOB. Let us stress here that Assumption 7.2 (resp. Assumption 7.1) is a condition under which we are able to run the proof for the discrete time optimization of the price (resp. volume) reversion models. From a practical point of view, it may seem too restrictive or unrealistic. However, these conditions are sufficient but not necessary. Looking at the proofs, we can observe that the optimal strategy $\boldsymbol{\xi}^\star$ given by Theorem 7.5 (resp. 7.3), if it is well-defined, is still a critical point of the cost function when Assumption 7.2 (resp. 7.1) does not hold. Of course, this is not enough to conclude that $\boldsymbol{\xi}^\star$ is a global minimizer, but indicates that $\boldsymbol{\xi}^\star$ is a good candidate for that. Another way to avoid Assumption 7.2 (resp. 7.1) is to work in continuous time and check directly the hypothesis of Theorem 7.6 (resp. 7.4).

Theorem 7.5. *Under Assumption 7.2, the cost function $C^P(\boldsymbol{\xi}, \boldsymbol{t})$ is nonnegative, and there is a unique optimal execution strategy $\boldsymbol{\xi}^\star$ that minimizes C^P over $\{\boldsymbol{\xi} \in \mathbb{R}^{N+1}, \sum_{i=0}^N \xi_i = -\mathrm{x}\}$. This strategy is given as follows. The following equation*

$$\sum_{i=1}^{N} \lambda(t_{i-1}) \left[F\left(\frac{h_{P,i}^{-1}(\nu)}{a_i} \right) - \frac{\lambda(t_i)}{\lambda(t_{i-1})} F(h_{P,i}^{-1}(\nu)) \right] + \lambda(t_N) F(\nu) = -\mathrm{x}$$

has a unique solution $\nu \in \mathbb{R}$, and

$$\begin{aligned}
\xi_0^\star &= \lambda(t_0) F\left(\frac{h_{P,1}^{-1}(\nu)}{a_1} \right), \\
\xi_i^\star &= \lambda(t_i) \left[F\left(\frac{h_{P,i+1}^{-1}(\nu)}{a_{i+1}} \right) - F(h_{P,i}^{-1}(\nu)) \right],\ 1 \le i \le N-1, \\
\xi_N^\star &= \lambda(t_N)[F(\nu) - F(h_{P,N}^{-1}(\nu))].
\end{aligned}$$

The first and the last trade have the same sign as $-\mathrm{x}$.

We now state the corresponding result in continuous time and set:

$$x \in \mathbb{R},\ h_{P,t}(x) = x \left[1 + \frac{\rho_t}{\rho_t \left(1 + \frac{x f'(x)}{f(x)} \right) - \eta_t} \right]. \tag{7.37}$$

Theorem 7.6. *Let $f \in \mathcal{C}^2(\mathbb{R})$. We assume that one of the two following conditions holds.*

1. *For $t \in [0,T]$, $\rho_t \left(1 + \frac{x f'(x)}{f(x)} \right) - \eta_t > 0$ for any $x \in \mathbb{R}$ and $h_{P,t}$ is bijective on $\mathbb{R}$, such that $h'_{P,t}(x) > 0$, dx-a.e.*
2. *For $t \in [0,T]$, $\rho_t \left(1 + \frac{x f'(x)}{f(x)} \right) - \eta_t < 0$ and $\rho_t \left(2 + \frac{x f'(x)}{f(x)} \right) - \eta_t > 0$ for any $x \in \mathbb{R}$, and $h_{P,t}$ is bijective on $\mathbb{R}$, such that $h'_{P,t}(x) < 0$, dx-a.e.*

Then, the cost function $C^P(X)$ is nonnegative, and there is a unique optimal admissible strategy $X^\star$ that minimizes C^P. This strategy is given as follows. The equation

$$\int_0^T \lambda(t) [\rho_t h_{P,t}^{-1}(\nu) f(h_{P,t}^{-1}(\nu)) - \eta_t F(h_{P,t}^{-1}(\nu))] dt + \lambda(T) F(\nu) = -\mathrm{x} \tag{7.38}$$

has a unique solution $\nu \in \mathbb{R}$ and we set $\zeta_t = h_{P,t}^{-1}(\nu)$. The strategy $dX_t^\star = \xi_0^\star \delta_0(dt) + \xi_t^\star dt + \xi_T^\star \delta_T(dt)$ with

$$\xi_0^\star = \lambda(0)F(\zeta_0),$$
$$\xi_t^\star = \lambda(t)f(\zeta_t)\left[\frac{d\zeta_t}{dt} + \rho_t\zeta_t\right],$$
$$\xi_T^\star = \lambda(T)(F(\nu) - F(\zeta_T)),$$

is optimal. The initial trade $\xi_0^\star$ has the same sign as $-\mathbb{x}$.

In particular, there is no PMS in model P as soon as Assumptions (i) or (ii) hold. Conversely, let us assume that $\rho_{t_1}\left(2 + \frac{xf'(x)}{f(x)}\right) - \eta_{t_1} < 0$, when x belongs to a neighbourhood of 0 for some $t_1 \geq 0$. This is equivalent to assume that $2\rho_{t_1} - \eta_{t_1} < 0$ for some $t_1 \geq 0$. Then, we set $\mathbf{t} = (0, t_1, t_2)$ with $t_2 > t_1$, and consider that the large trader buys $x > 0$ at time t_1 and sells x at time t_2. The cost of such a round trip is

$$\begin{aligned}
&C^P((0, x, -x), \boldsymbol{t}) \\
&= \lambda(t_1)G\left(\frac{x}{\lambda(t_1)}\right) + \lambda(t_2)\left[G\left(F\left(e^{-\int_{t_1}^{t_2}\rho_s ds}F^{-1}\left(\frac{x}{\lambda(t_2)}\right)\right) - \frac{x}{\lambda(t_2)}\right) - \tilde{F}\left(e^{-\int_{t_1}^{t_2}\rho_s ds}F^{-1}\left(\frac{x}{\lambda(t_2)}\right)\right)\right] \\
&= \lambda(t_1)\left[-\eta_{t_1}\tilde{F}\left(F^{-1}\left(\frac{x}{\lambda(t_1)}\right)\right) + \rho_{t_1}F^{-1}\left(\frac{x}{\lambda(t_1)}\right)^2 f\left(F^{-1}\left(\frac{x}{\lambda(t_1)}\right)\right)\right](t_2 - t_1) + o(t_2 - t_1).
\end{aligned}$$

The derivative of $x \mapsto -\eta_{t_1}\tilde{F}(x) + \rho_{t_1}x^2 f(x)$ is $xf(x)\left(\rho_{t_1}\left(2 + \frac{xf'(x)}{f(x)}\right) - \eta_{t_1}\right)$ and has the opposite sign of x near 0. Thus, $C^P((0, x, -x), \boldsymbol{t})$ is negative when t_2 is close to t_1 and x is small enough, which gives a PMS.

Corollary 7.6. *Let $f \in \mathcal{C}^2(\mathbb{R})$. Under Assumption 7.2, the function $h_{P,t}$ is $\mathcal{C}^1(\mathbb{R})$, bijective on $\mathbb{R}$ and such that $h'_{P,t} > 0$. Thus, the result of Theorem 7.6 holds and the last trade $\xi_T^\star$ has the same sign as $-\mathbb{x}$.*

Besides, if (7.30) *also holds, $\xi_t^\star$ has the same sign as $-\mathbb{x}$ for any $0 < t < T$, which rules out TTPM.*

As for model V, we consider now the case of a power-law shape $f(x) = |x|^\gamma$. We can apply the results of Theorem 7.6 in this case. We can also notice from (7.9) that $dE_t = (\eta_t - \rho_t(1+\gamma))E_t dt$. Therefore, model P with resilience ρ_t is the same as model V with resilience $\tilde{\rho}_t = \rho_t(1+\gamma) - \eta_t$.

Corollary 7.7. *When $f(x) = |x|^\gamma$, model P does not admit PMS if, and only if*

$$\forall t \geq 0,\ \ \rho_t(2+\gamma) - \eta_t \geq 0.$$

It does not admit transaction-triggered price manipulation if and only if

$$\forall t \geq 0,\ \ \rho_t(1+\gamma) - \eta_t \geq 0,\ \ \textit{and}\ \ \left(\frac{\rho_t(1+\gamma) - \eta_t}{\rho_t(2+\gamma) - \eta_t}\right)' + \rho_t\left(\frac{\rho_t(1+\gamma) - \eta_t}{\rho_t(2+\gamma) + \eta_t}\right) \geq 0.$$

7.2.3 Numerical results

In this paragraph, we would like to briefly investigate the impact of the different parameters on the optimal strategy. As already mentioned in this paper, it has been observed in the literature that the price impact of an order is proportional to a power of its volume (see e.g. Tóth et. al [BDdL$^+$11]), this power being close to $1/2$. This leads us to consider the power law shape $f(x) = |x|^\gamma$. In this model, the shift on the price D_t is proportional to $E_t^{\frac{1}{\gamma+1}}$, which indicates the choice of $\gamma \approx 1$. Last, we consider the volume impact reversion model. As said above, this choice is not restrictive since for a power law shape function it is equivalent to the price impact reversion model, up to a change of the resilience function.

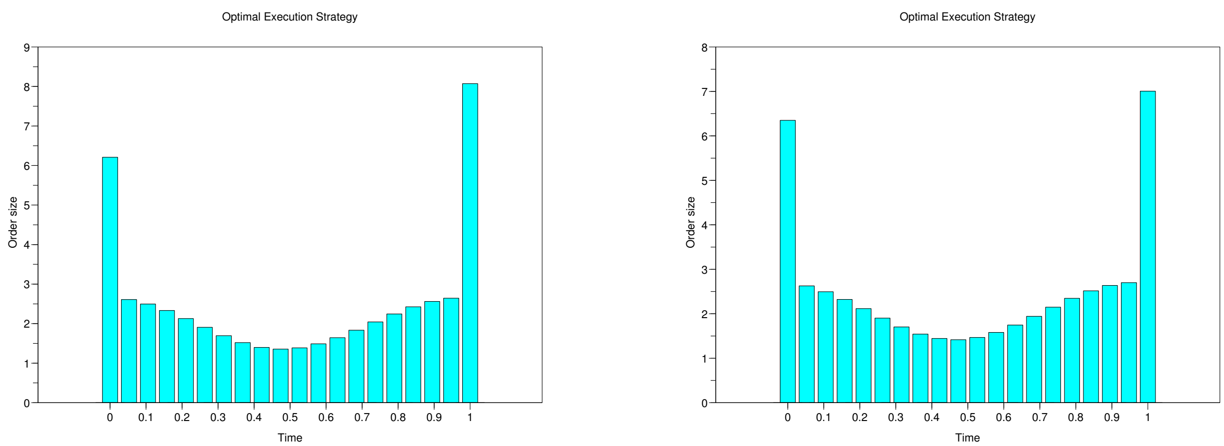


Fig. 7.3. Optimal execution strategy to buy 50 shares on a regular time grid, with $N = 20$, $\rho = 10$, $\lambda(t) = 4 + \cos(2\pi t)$ for $\gamma = 7/3$ (left) or $\gamma = 3/7$ (right).

We first examine how the shape function modifies the optimal strategy. We have already seen in Figure 7.2 that this function may significantly change the optimal trades, as one could expect. However, this Figure consider two rather different values of γ with parameters implying TTPM. Now, we consider other parameters that exclude TTPM since we deem it should be the case for in practice. We still consider a one day (i.e. one period) deadline. As recalled above, it is empirically observed that the price impact of a trade is proportional to a power of the trade size, and this power is typically close to $1/2$. In Figure (7.3), we have plotted the optimal strategy for $\gamma = 7/3$ and $\gamma = 3/7$, which corresponds respectively to a power $1/(\gamma + 1)$ equal to 0.3 and 0.7. The optimal strategy for $\gamma = 1$ (i.e. a square-root impact) is plotted in the upper right part of Figure (7.4). Qualitatively, there is no striking difference between these optimal strategies that do not seem very sensitive to γ. The main change is the size of the last trade that increases when γ gets larger.

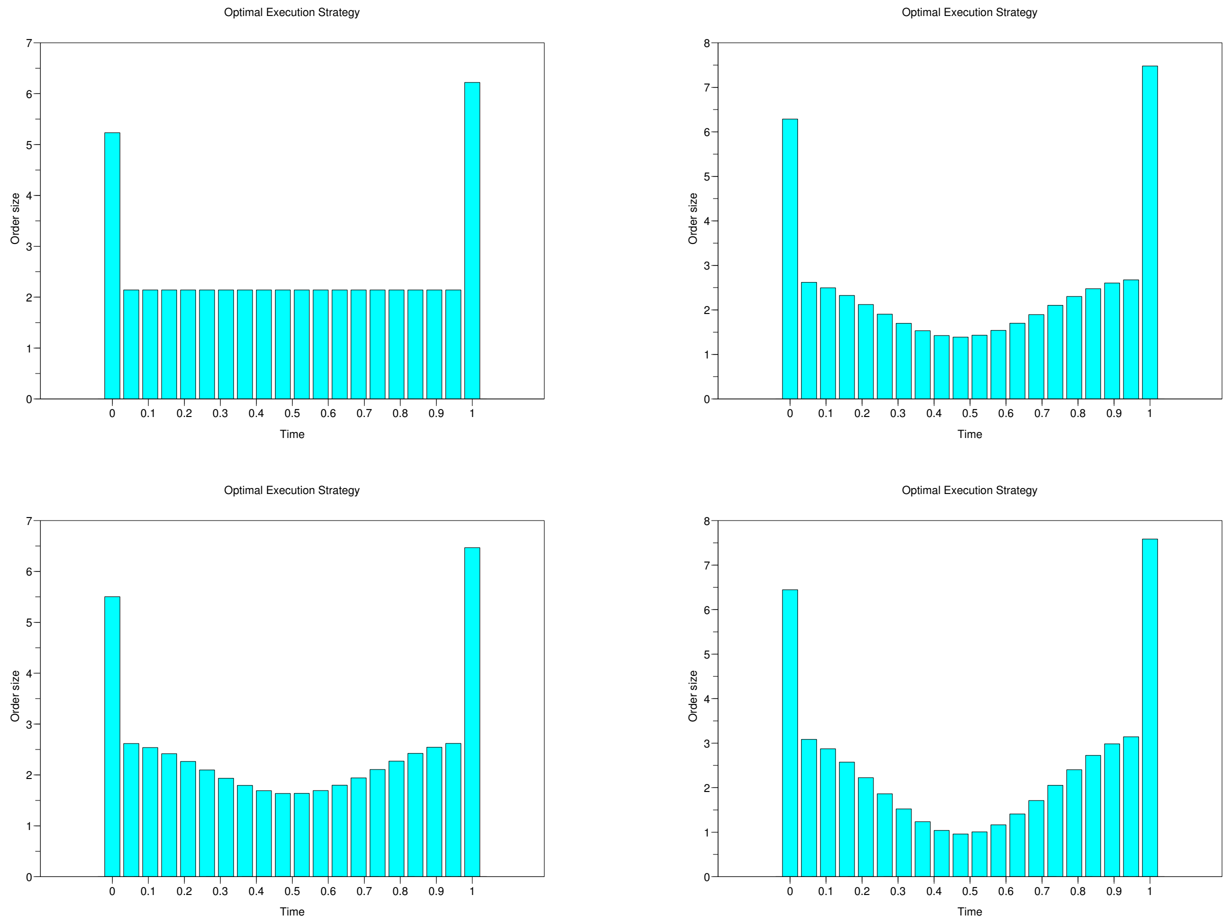


Fig. 7.4. Optimal execution strategy to buy 50 shares on a regular time grid, with $N = 20$, $\lambda(t) = 4 + \varepsilon_\lambda \cos(2\pi t)$, $\rho(t) = 10 + 2.5\varepsilon_\rho \cos(2\pi t)$ and $\gamma = 1$ for $(\varepsilon_\lambda, \varepsilon_\rho)$ equal to $(0, 0)$ (upper left), $(1, 0)$ (upper right), $(0, 1)$ (lower left) or $(1, 1)$ (lower right).

In Figure 7.4 we focus on how is modified the optimal strategy when we consider constant or time dependent functions λ and ρ. Let us recall that λ models the liquidity available on the market while ρ models its resilience. It is known from Alfonsi, Fruth and Schied [AFS10] that the optimal strategy is made with intermediate trades of the same size when ρ and λ are both constant. This is again observed in the upper left graphic of Figure 7.4. Then, we consider in the upper right and the lower left graphics the case where λ depends on t and ρ is constant, and conversely. To compare, we have considered each time the same type of time-dependency, which is proportional to $4 + \cos(2\pi t)$. Qualitatively, this two graphics are rather similar. As one could expect, the size of the intermediate trades follows the trend of the market liquidity or resilience. Looking closer, this impact on the optimal strategy is more important for the liquidity oscillations than for the resilience. Last, when both liquidity and resilience depend on the time (lower right graphic), these effects add up. The intermediate trades change significantly along the time. Also, the first and last trades are larger, because they coincides to the upper value of λ and ρ. Thus, we see that the evolution of λ and ρ have a clear impact on the optimal strategy: this strategy follows the trend given by λ and ρ. This behaviour is very close to the VWAP (Volume Weighted Average Price) strategy which is widely used by practitioners.

7.3 Proofs

7.3.1 The block shape case

Proof of Theorem 7.1: The quadratic form (7.16) is given by $C^V(\boldsymbol{\xi}, \boldsymbol{t}) = \frac{1}{2}\boldsymbol{\xi}^T M^V \boldsymbol{\xi}$, with $M^V_{i,j} = \frac{\exp\left(-\left|\int_{t_i}^{t_j} \rho_s ds\right|\right)}{\lambda(t_i \vee t_j)}$, $0 \leq i, j \leq N$. Let us assume that $a_i \tilde{a}_i < 1, \forall i \in \{1, \dots, N\}$. Then, we can define the following vectors:

$$\mathbf{y}_0 = \frac{\mathbf{e}_0}{\sqrt{\lambda(t_0)}}, \ \mathbf{y}_i = \tilde{a}_i \mathbf{y}_{i-1} + \frac{\mathbf{e}_i}{\sqrt{\lambda(t_i)}} \sqrt{1 - a_i \tilde{a}_i}, \ 1 \leq i \leq N$$

where $\mathbf{e}_0 \dots \mathbf{e}_N$ denote the canonical basis of $\mathbb{R}^{N+1}$. We have $M^V_{ij} = \mathbf{y}_i^T \mathbf{y}_j$. We introduce Y the upper triangular matrix with columns $\mathbf{y}_0, \dots, \mathbf{y}_N$. By assumption, it is invertible and so is $M = Y^T Y$. Conversely, if M^V is positive definite, the minors

$$\det((M^V_{i,j})_{0 \leq i,j \leq n}) = \frac{1}{\lambda(t_0)} \prod_{i=1}^{n} \frac{1}{\lambda(t_i)} (1 - a_i \tilde{a}_i), \ 1 \leq n \leq N$$

are positive, which gives (7.17).

Let us turn to the optimization problem. One has to minimize $C^V(\boldsymbol{\xi}, \boldsymbol{t})$ under the linear constraint $\sum_{i=0}^{N} \xi_i = -\mathbb{x}$, which gives

$$\boldsymbol{\xi}^\star = -\frac{\mathbb{x}}{\mathbf{1}^T \left(M^V\right)^{-1} \mathbf{1}} \left(M^V\right)^{-1} \mathbf{1}, \tag{7.39}$$

where $\mathbf{1} \in \mathbb{R}^{N+1}$ is a vector of ones. Since Y is upper triangular, it can be easily inverted and we can calculate explicitly $\left(M^V\right)^{-1} \mathbf{1}$ and get (7.18). We do not detail these calculations since the result is anyway a consequence of Theorem 7.3. □

7.3.2 General limit order book shape with model *V*

Let us introduce some notations. For the time grid $\boldsymbol{t}$ given by (7.14), we introduce the following quantities:

$$\alpha_k := \int_{t_{k-1}}^{t_k} \rho_s ds,\ k = 1, \dots, N. \tag{7.40}$$

We can write the cost function (7.13) as follows

$$C^V(\boldsymbol{\xi}, \boldsymbol{t}) = \sum_{n=0}^{N} \lambda(t_n) \left[G\left(\frac{E_n + \xi_n}{\lambda(t_n)} \right) - G\left(\frac{E_n}{\lambda(t_n)} \right) \right], \tag{7.41}$$

where we use the following notations (observe that $E_n = a_n(E_{n-1} + \xi_{n-1})$)

$$E_0 = 0,\ E_n = \sum_{i=0}^{n-1} \xi_i e^{-\sum_{k=i+1}^{n} \alpha_k},\ 1 \le n \le N.$$

Lemma 7.1. *We have* $\frac{\partial C^V}{\partial \xi_N} = F^{-1}\left(\frac{E_N + \xi_N}{\lambda(t_N)}\right)$ *and, for* $i = 0, \dots, N-1$,

$$\frac{\partial C^V}{\partial \xi_i} - a_{i+1} \frac{\partial C^V}{\partial \xi_{i+1}} = F^{-1}\left(\frac{E_i + \xi_i}{\lambda(t_i)} \right) - a_{i+1} F^{-1}\left(\frac{E_{i+1}}{\lambda(t_{i+1})} \right). \tag{7.42}$$

Proof. Let us first observe that $\frac{\partial E_n}{\partial \xi_i} = 0$, if $i \ge n$, and $\frac{\partial E_n}{\partial \xi_i} = e^{-\sum_{k=i+1}^{n} \alpha_k}$ if $i < n$. Thus, we get by using that $G' = F^{-1}$:

$$\begin{aligned}
\frac{\partial C^V}{\partial \xi_i} &= F^{-1}\left(\frac{E_i + \xi_i}{\lambda(t_i)} \right) + \sum_{n=i+1}^{N} e^{-\sum_{k=i+1}^{n} \alpha_k} \left(F^{-1}\left(\frac{E_n + \xi_n}{\lambda(t_n)} \right) - F^{-1}\left(\frac{E_n}{\lambda(t_n)} \right) \right) \\
&= F^{-1}\left(\frac{E_i + \xi_i}{\lambda(t_i)} \right) - e^{-\alpha_{i+1}} F^{-1}\left(\frac{E_{i+1}}{\lambda(t_{i+1})} \right) \\
&\quad + e^{-\alpha_{i+1}} \left[F^{-1}\left(\frac{E_{i+1} + \xi_{i+1}}{\lambda(t_{i+1})} \right) + \sum_{n=i+2}^{N} e^{-\sum_{k=i+2}^{n} \alpha_k} \left(F^{-1}\left(\frac{E_n + \xi_n}{\lambda(t_n)} \right) - F^{-1}\left(\frac{E_n}{\lambda(t_n)} \right) \right) \right] \\
&= F^{-1}\left(\frac{E_i + \xi_i}{\lambda(t_i)} \right) - a_{i+1} F^{-1}\left(\frac{E_{i+1}}{\lambda(t_{i+1})} \right) + a_{i+1} \frac{\partial C^V}{\partial \xi_{i+1}}.
\end{aligned}$$

Lemma 7.2. *Under Assumption 7.1, we obtain the next conclusions.*

1. *For* $i \in \{1, \dots, N\}$, *the function* $h_{V,i}$ *defined in* (7.32) *is an increasing bijection on* $\mathbb{R}$ *that satisfies* $\operatorname{sgn}(x) h_{V,i}(x) \ge \frac{1 - a_i \tilde{a}_i}{1 - a_i} F^{-1}(x)$.
2. *If* (7.33) *holds, then we have* $\operatorname{sgn}(x) h_{V,i+1}^{-1}(x) \ge \operatorname{sgn}(x) \tilde{a}_i h_{V,i}^{-1}(x)$ *for* $i \in \{1, \dots, N-1\}$.
3. $\operatorname{sgn}(x) F(x) \ge \operatorname{sgn}(x) \tilde{a}_N h_{V,N}^{-1}(x)$.

Proof. 1. Since the resilience ρ_t is positive, we have $0 < a_i < 1$, and $\tilde{a}_i \le 1$ since $\rho_t + \eta_t \ge 0$ by Assumption 7.1. We then get

$$\frac{\partial h_{V,i}(x)}{\partial x} = \frac{1}{1 - a_i} \left[\frac{1}{f(F^{-1}(x))} - \frac{a_i \tilde{a}_i}{f(F^{-1}(\tilde{a}_i x))} \right] \ge \frac{1 - a_i \tilde{a}_i}{1 - a_i} \frac{1}{f(F^{-1}(x))} > 0$$

because f is nondecreasing on $\mathbb{R}_-$ and nonincreasing on $\mathbb{R}_+$, and F^{-1} is increasing.

2. We set $\hat{f}(x) = (F^{-1})'(x) = 1/f(F^{-1}(x))$: this function is positive, nonincreasing on $\mathbb{R}_-$ and nondecreasing on $\mathbb{R}_+$. Let $\nu \geq 0$ and $y = h_{V,i+1}^{-1}(\nu)$. We note that $y \geq 0$ because $h_{V,i+1}(0) = 0$ and $h_{V,i+1}$ is increasing by the first point of this lemma. Thus, we have that

$$\begin{aligned}\nu &= \frac{F^{-1}(y) - a_{i+1}F^{-1}(\tilde{a}_{i+1}y)}{1-a_{i+1}} \\ &= F^{-1}(\tilde{a}_{i+1}y) + \frac{F^{-1}(y) - F^{-1}(\tilde{a}_{i+1}y)}{1-a_{i+1}} \\ &= F^{-1}(\tilde{a}_{i+1}y) + \frac{1}{1-a_{i+1}}\int_{\tilde{a}_{i+1}y}^{y} \hat{f}(\xi)d\xi \leq F^{-1}(y) + \frac{1-\tilde{a}_{i+1}}{1-a_{i+1}} y\hat{f}(y) =: g_{i+1}(y)\end{aligned}$$

Hence, we obtain that g_{i+1} is increasing on $\mathbb{R}$ and then, $y \geq g_{i+1}^{-1}(\nu)$. Let $z = \tilde{a}_i h_{V,i}^{-1}(\nu) \geq 0$. We have:

$$\begin{aligned}\nu &= \frac{F^{-1}\left(\frac{z}{\tilde{a}_i}\right) - a_i F^{-1}(z)}{1-a_i} \\ &= F^{-1}(z) + \frac{F^{-1}\left(\frac{z}{\tilde{a}_i}\right) - F^{-1}(z)}{1-a_i} \\ &= F^{-1}(z) + \frac{1}{1-a_i}\int_{z}^{\frac{z}{\tilde{a}_i}} \hat{f}(\xi)d\xi \geq F^{-1}(z) + \frac{\left(\frac{1}{\tilde{a}_i} - 1\right)}{1-a_i} z\hat{f}(z) =: \bar{g}_i(z)\end{aligned}$$

Therefore, if (7.33) holds, we get that $g_{i+1}(x) \leq \bar{g}_i(x)$ for all $x \geq 0$. Then, we have $g_{i+1}^{-1}(x) \geq g_i^{-1}(x)$, and therefore

$$y \geq g_{i+1}^{-1}(\nu) \geq g_i^{-1}(\nu) \geq z.$$

The same arguments for $\nu \leq 0$ give $y \leq g_{i+1}^{-1}(\nu) \leq g_i^{-1}(\nu) \leq z$.

3. Using the above definition, we have $\operatorname{sgn}(x)\bar{g}_N(x) \geq \operatorname{sgn}(x)F^{-1}(x)$, and therefore we get

$$\operatorname{sgn}(\nu)F(\nu) \geq \operatorname{sgn}(\nu)\bar{g}_N^{-1}(\nu) \geq \operatorname{sgn}(\nu)z = \operatorname{sgn}(\nu)\tilde{a}_N h_{V,N}^{-1}(\nu)\,.$$

Lemma 7.3. *Let $a \in (0,1)$ and $b > 0$ such that $ab \leq 1$. We have $G(x) - \frac{1}{b}G(abx) \geq 0$ for $x \in \mathbb{R}$, and $G(x) - \frac{1}{b}G(abx) \underset{|x|\to+\infty}{\to} +\infty$.*

Proof. Since G is convex ($G' = F^{-1}$ is increasing) and $G(0) = 0$, $G(abx) \leq abG(x)$. If $b > 1$, we then have $G(x) - \frac{1}{b}G(abx) \geq G(x)(1-a)$ which gives the result. If $b \leq 1$, we have

$$\begin{aligned}G(x) - \frac{1}{b}G(abx) &= \int_0^x F^{-1}(u)du - \frac{1}{b}\int_0^{abx} F^{-1}(u)du = \int_0^x F^{-1}(u)du - \int_0^{ax} F^{-1}(bv)dv \\ &= \int_{ax}^x F^{-1}(u)du + \int_0^{ax}\Big(F^{-1}(u) - F^{-1}(bu)\Big)\,du \geq |x|(1-a)F^{-1}(|ax|) \underset{|x|\to+\infty}{\to} \infty.\end{aligned}$$

Proof of Theorem 7.3: We rewrite the cost function (7.41) to minimize as follows:

$$\begin{aligned}
C^V(\boldsymbol{\xi},\boldsymbol{t}) &= \sum_{n=0}^{N} \lambda(t_n)\left[G\left(\frac{E_n+\xi_n}{\lambda(t_n)}\right) - G\left(\frac{E_n}{\lambda(t_n)}\right)\right] \\
&= \lambda(t_N) G\left(\frac{\sum_{i=0}^{N} \xi_i e^{-\sum_{k=i+1}^{N}\alpha_k}}{\lambda(t_N)}\right) - \lambda(0)G(0) \\
&+ \sum_{n=0}^{N-1}\left[\lambda(t_n)G\left(\frac{\sum_{i=0}^{n}\xi_i e^{-\sum_{k=i+1}^{n}\alpha_k}}{\lambda(t_n)}\right) - \lambda(t_{n+1})G\left(\frac{e^{-\alpha_{n+1}}\sum_{i=0}^{n}\xi_i e^{-\sum_{k=i+1}^{n}\alpha_k}}{\lambda(t_{n+1})}\right)\right]
\end{aligned}$$

We define the linear map $T:\mathbb{R}^{N+1}\to\mathbb{R}^{N+1}$ by $(T\xi)_n = \frac{\sum_{i=0}^{n}\xi_i e^{-\sum_{k=i+1}^{n}\alpha_k}}{\lambda(t_n)}$, so that

$$C^V(\boldsymbol{\xi},\boldsymbol{t}) = \lambda(t_N)G((T\xi)_N) + \sum_{n=0}^{N-1}\left[\lambda(t_n)G((T\xi)_n) - \lambda(t_{n+1})G\left(\tilde{a}_{n+1}(T\xi)_n\right)\right]. \qquad (7.43)$$

Let us observe that T is a linear bijection. By Lemma 7.3 we get that $C^V(\boldsymbol{\xi},\boldsymbol{t})\geq 0$ and $C^V(\boldsymbol{\xi},\boldsymbol{t}) \underset{|\boldsymbol{\xi}|\to+\infty}{\to} +\infty$, which gives the existence of a minimizer $\boldsymbol{\xi}^\star$ over $\boldsymbol{\xi}$, s.t. $\sum_{i=0}^{N}\xi_i = -\mathbb{x}$. Thus, by using (7.42), there must be a Lagrange multiplier ν such that

$$\nu = h_{V,i+1}\left(\frac{E_i+\xi_i^\star}{\lambda(t_i)}\right),\ i=0\ldots N-1, \text{ and } \nu = F^{-1}\left(\frac{E_N+\xi_N^\star}{\lambda(t_N)}\right). \qquad (7.44)$$

We have $\frac{E_i+\xi_i^\star}{\lambda(t_i)} = h_{V,i+1}^{-1}(\nu)$ and then $E_{i+1} = \lambda(t_i)a_{i+1}h_{V,i+1}^{-1}(\nu)$, for $0\leq i\leq N-1$. Thus, we get

$$\begin{aligned}
\xi_0^\star &= \lambda(t_0)h_{V,1}^{-1}(\nu), \\
\xi_i^\star &= \lambda(t_i)h_{V,i+1}^{-1}(\nu) - \lambda(t_{i-1})a_i h_{V,i}^{-1}(\nu),\ 1\leq i\leq N-1, \\
\xi_N^\star &= F(\nu)\lambda(t_N) - \lambda(t_{N-1})a_N h_{V,N}^{-1}(\nu)
\end{aligned}$$

Furthermore, we note that

$$\sum_{i=0}^{N}\xi_i^\star = -\mathbb{x} = \lambda(t_0)(1-a_1)h_{V,1}^{-1}(\nu) + \ldots + \lambda(t_{N-1})(1-a_N)h_{V,N}^{-1}(\nu) + F(\nu)\lambda(t_N).$$

By Lemma 7.2 The right side is an increasing bijection on $\mathbb{R}$, and we deduce that there is only one $\nu\in\mathbb{R}$ which satisfies the above equation. This give the uniqueness of the minimizer $\boldsymbol{\xi}^\star$. Moreover, the functions F^{-1} and $h_{V,i}$ vanish in 0, and ν has the same sign as $-\mathbb{x}$, which gives that $\xi_0^\star$ and $\xi_N^\star$ have the same sign as $-\mathbb{x}$ by Lemma 7.2. Besides, if (7.33) holds, the trades $\xi_i^\star$ have also the same sign as $-\mathbb{x}$. □

Let us now prepare the proof of Theorem 7.4 and assume that $h_{V,t}$ is bijective increasing. We introduce for $0\leq t\leq T$,

$$C^V(t,T,E_t,X_t) = \lambda(t)\left[G(\zeta_t) - G\left(\frac{E_t}{\lambda(t)}\right)\right] + \int_t^T F^{-1}(\zeta_u)\xi_u du + \lambda(T)[G(F(\nu)) - G(\zeta_T)], \quad (7.45)$$

where

$$\nu \in \mathbb{R}, s.t. - E_t + \int_t^T \lambda(u)\rho_u h_{V,u}^{-1}(\nu)du + \lambda(T)F(\nu) = -X_t, \quad (7.46)$$

$$\zeta_u = h_{V,u}^{-1}(\nu),\ \xi_u = \lambda(u)[\frac{d\zeta_u}{du} + (\rho_u + \eta_u)\zeta_u]. \quad (7.47)$$

Let us observe that $\nu \mapsto \int_t^T \lambda(u)\rho_u h_{V,u}^{-1}(\nu)du + \lambda(T)F(\nu)$ is increasing and bijective on $\mathbb{R}$, and (7.46) admits a unique solution. The function $C^V(t,T,E_t,X_t)$ denotes the minimal cost to liquidate X_t shares on the time interval $[t,T]$ given the current state E_t. In particular, we observe that

$$C^V(T,T,E_T,X_T) = \lambda(T)\left[G\left(\frac{E_T - X_T}{\lambda(T)}\right) - G\left(\frac{E_T}{\lambda(T)}\right)\right],$$

which is the cost of selling X_T shares at time T. Besides, an integration by parts gives that

$$C^V(t,T,E_t,X_t) = -\lambda(t)G\left(\frac{E_t}{\lambda(t)}\right) + \int_t^T \lambda(u)\left[(\rho_u+\eta_u)F^{-1}(\zeta_u)\zeta_u - \eta_u G(\zeta_u)\right]du + \lambda(T)G(F(\nu)). \quad (7.48)$$

The function $\zeta \mapsto (\rho_u+\eta_u)F^{-1}(\zeta)\zeta - \eta_u G(\zeta)$ is nonnegative since it vanishes at 0, and its derivative is equal to $\rho_u h_{V,u}(\zeta)$ that has the same sign as ζ. Since $G \geq 0$, we get:

$$C^V(0,T,0,\mathbb{x}) \geq 0. \quad (7.49)$$

Formula (7.45) can be guessed by simple but tedious calculations: one has to consider the associated discrete problem on a regular time-grid and then let the time-step going to zero. We do not present these calculations here since we will prove directly by a verification argument that this is indeed the minimal cost.

Proof of Theorem 7.4: Let $(X_t, 0 \leq t \leq T+)$ denote an admissible strategy that liquidates $\mathbb{x}$. We consider $(E_t, 0 \leq t \leq T+)$ the solution of $dE_t = dX_t - \rho_t E_t dt$, ν_t the solution of (7.46) and $\zeta_t = h_{V,t}^{-1}(\nu_t)$. We set

$$C_t = \int_0^t F^{-1}\left(\frac{E_s}{\lambda(s)}\right)dX_s^c + \sum_{0\leq s<t}\lambda(s)\left[G\left(\frac{E_s+\Delta X_s}{\lambda(s)}\right) - G\left(\frac{E_s}{\lambda(s)}\right)\right] + C^V(t,T,E_t,X_t).$$

Let us observe that $C_T = C^V(X)$ and $C_0 = C^V(0,T,0,\mathbb{x})$. We are going to show that $dC_t \geq 0$, and that $dC_t = 0$ holds only for $X^\star$. This will in particular show that $C^V(X) \geq 0$ from (7.49).

Let us first consider the case of a jump $\Delta X_t > 0$. Then, we have

$$\Delta C_t = \lambda(t)\left[G\left(\frac{E_t+\Delta X_t}{\lambda(t)}\right) - G\left(\frac{E_t}{\lambda(t)}\right)\right] + C^V(t+,T,E_{t+},X_{t+}) - C^V(t,T,E_t,X_t).$$

Since $\Delta E_t = \Delta X_t$, the solution ν_t of (7.46) is also the solution of $-E_{t+} + \int_t^T \lambda(u)\rho_u h_{V,u}^{-1}(\nu_t)du + \lambda(T)F(\nu_t) = -X_{t+}$, and then $\Delta C_t = 0$. Now, let us calculate dC_t. We set

$$\tilde{C}(t,T,E_t,X_t,v) = \lambda(T)G(F(v)) - \lambda(t)G\left(\frac{E_t}{\lambda(t)}\right) + \int_t^T \lambda(u)\left[(\rho_u+\eta_u)F^{-1}(h_{V,u}^{-1}(v))h_{V,u}^{-1}(v) - \eta_u G(h_{V,u}^{-1}(v))\right]du.$$

Then, we have from (7.48):

$$dC_t = F^{-1}\left(\frac{E_t}{\lambda(t)}\right)dX_t^c - \lambda'(t)G\left(\frac{E_t}{\lambda(t)}\right)dt - F^{-1}\left(\frac{E_t}{\lambda(t)}\right)(dX_t^c - (\rho_t+\eta_t)E_t dt) - \lambda(t)(\rho_t+\eta_t)F^{-1}(\zeta_t)\zeta_t dt + \lambda'(t)G(\zeta_t)dt + \frac{\partial \tilde{C}}{\partial v}(t,T,E_t,X_t,\nu_t)d\nu_t.$$

Since $\left[\lambda(T)f(\nu_t) + \int_t^T \lambda(u)\rho_u (h_{V,u}^{-1})'(\nu_t)du\right]d\nu_t - \lambda(t)\rho_t h_{V,t}^{-1}(\nu_t)dt = d(E_t - X_t) = -\rho_t E_t dt$ and

$$\partial_v \tilde{C}(t,T,E_t,X_t,v) = \lambda(T)vf(v) + \int_t^T \lambda(u)\rho_u (h_{V,u}^{-1})'(v)\left[F^{-1}(h_{V,u}^{-1}(v)) + \frac{\rho_u+\eta_u}{\rho_u}\frac{h_{V,u}^{-1}(v)}{f(h_{V,u}^{-1}(v))}\right]du$$
$$= v\left[\lambda(T)f(v) + \int_t^T \lambda(u)\rho_u (h_{V,u}^{-1})'(v)du\right],$$

we finally get

$$dC_t = \lambda(t)\left[(\rho_t+\eta_t)\left(\frac{E_t}{\lambda(t)}F^{-1}\left(\frac{E_t}{\lambda(t)}\right) - \zeta_t F^{-1}(\zeta_t)\right) + \eta_t\left(G(\zeta_t) - G\left(\frac{E_t}{\lambda(t)}\right)\right) + \rho_t h_{V,t}(\zeta_t)\left(\zeta_t - \frac{E_t}{\lambda(t)}\right)\right]dt$$
$$:= \lambda(t)\psi_t(\zeta_t)dt. \tag{7.50}$$

We have $\psi_t'(\zeta) = -(\rho_t+\eta_t)\left(F^{-1}(\zeta) + \frac{\zeta}{f(F^{-1}(\zeta))}\right) + \eta_t F^{-1}(\zeta) + \rho_t h_{V,t}(\zeta) + \rho_t h_{V,t}'(\zeta)(\zeta - \frac{E_t}{\lambda(t)}) = \rho_t h_{V,t}'(\zeta)(\zeta - \frac{E_t}{\lambda(t)})$. Since $h_{V,t}' > 0$, ψ_t vanishes at $\zeta = \frac{E_t}{\lambda(t)}$, and is positive for $\zeta \neq \frac{E_t}{\lambda(t)}$.

Thus, if X is an optimal strategy, we necessarily have $\zeta_t = \frac{E_t}{\lambda(t)}$, dt-a.e. Then, we get by differentiating $\left[X_t - E_t + \int_t^T \lambda(u)\rho_u h_{V,u}^{-1}(\nu_t)du + \lambda(T)F(\nu_t)\right] = 0$ that $\left[\int_t^T \lambda(u)\rho_u (h_{V,u}^{-1})'(\nu_t)du + \lambda(T)f(\nu_t)\right]d\nu_t = 0$, which gives $d\nu_t = 0$ since $(h_{V,u}^{-1})' > 0$ and $f > 0$. Thus, we get that $\nu_t = \nu$ where ν is the solution of (7.35). In particular, we get $\Delta X_0 = E_{0+} = \lambda(0)h_{V,0}^{-1}(0) = \Delta X_0^\star$ and then $X = X^\star$, which gives the uniqueness of the optimal strategy. Last we observe that ν has the same sign as $-\mathtt{x}$ and thus $\xi_0^\star$ has the same sign as $-\mathtt{x}$. □

Proof of Corollary 7.4: Since $\rho_t + \eta_t \geq 0$ and $xf'(F^{-1}(x)) \geq 0$ by Assumption 7.1, we have

$$h_{V,t}'(x) = \frac{\eta_t + 2\rho_t}{\rho_t}\frac{1}{f(F^{-1}(x))} - \frac{\eta_t+\rho_t}{\rho_t}\frac{xf'(F^{-1}(x))}{f(F^{-1}(x))^3} > 0.$$

Also, we have $\mathrm{sgn}(x)h_{V,t}(x) \geq \mathrm{sgn}(x)F^{-1}(x)$ and then $\mathrm{sgn}(x)h_{V,t}^{-1}(x) \leq \mathrm{sgn}(x)F(x)$, which gives that the last trade $\xi_T^\star$ has the same sign as $-\mathtt{x}$. Then, we have $\frac{d\zeta_t}{dt} = -\frac{1}{h_{V,t}'(\zeta_t)}\frac{dh_{V,t}}{dt}(\zeta_t)$ and thus

$$\xi_t^\star = \frac{\lambda(t)\zeta_t}{h'_{V,t}(\zeta_t)}\left[-\frac{d\,(\eta_t/\rho_t)}{dt}\frac{1}{f(F^{-1}(\zeta_t))} + (\rho_t+\eta_t)h'_{V,t}(\zeta_t)\right]$$
$$= \frac{\lambda(t)\zeta_t}{h'_{V,t}(\zeta_t)}\left[\frac{1}{\rho_t f(F^{-1}(\zeta_t))}\left(\frac{\rho_t'\eta_t-\rho_t\eta_t'}{\rho_t} + (\rho_t+\eta_t)(2\rho_t+\eta_t)\right) - \frac{(\eta_t+\rho_t)^2}{\rho_t}\frac{\zeta_t f'(\zeta_t)}{f(F^{-1}(\zeta_t))^3}\right]$$

is nonnegative if (7.22) holds since $h'_{V,t} > 0$ and $\zeta_t f'(\zeta_t) \geq 0$. □

Lemma 7.4. *We have* (7.33) $\implies$ (7.23) *if* $\rho_t + \eta_t \geq 0,\ t \geq 0$.

Proof. We have

$$(7.33) \Leftrightarrow \frac{1}{\tilde a_i}\frac{1-\tilde a_i}{1-a_i} \geq \frac{1-\tilde a_{i+1}}{1-a_{i+1}} \Leftrightarrow (1-a_{i+1}) - \tilde a_i\,(1-a_{i+1}) \geq \tilde a_i\,(1-a_i) - \tilde a_i \tilde a_{i+1}\,(1-a_i)$$
$$\Leftrightarrow \tilde a_{i+1}\,(1-a_i) + \frac{1}{\tilde a_i}(1-a_{i+1}) \geq 1-a_i+1-a_{i+1}.$$

Since $\tilde a_{i+1} \leq 1$, we get $1-a_i+1-a_{i+1} = 1-a_ia_{i+1}+(1-a_i)(1-a_{i+1}) \geq 1-a_ia_{i+1}+\tilde a_{i+1}(1-a_i)(1-a_{i+1})$. Thus, (7.33) implies that:

$$\tilde a_{i+1}\,(1-a_i) + \frac{1}{\tilde a_i}(1-a_{i+1}) \geq 1-a_ia_{i+1}+\tilde a_{i+1}(1-a_i)(1-a_{i+1})$$
$$\Leftrightarrow 1-\tilde a_i + a_ia_{i+1}\tilde a_i - a_i\tilde a_{i+1} \geq a_{i+1} - \tilde a_i\tilde a_{i+1}a_{i+1} + a_i\tilde a_i a_{i+1}\tilde a_{i+1} - \tilde a_{i+1}a_{i+1}$$
$$\Leftrightarrow (1-\tilde a_i)\,(1-a_{i+1}\tilde a_{i+1}) \geq a_{i+1}\,(1-\tilde a_{i+1})\,(1-a_i\tilde a_i) \Leftrightarrow (7.23).$$

7.3.3 General limit order book shape with model *P*

We first focus on discrete strategies on the time grid $\boldsymbol{t}$ such as (7.14). We introduce the following shorthand notation $D_n = D_{t_n}$ for $0 \leq n \leq N$ and have

$$D_0 = 0,\ D_n = a_n F^{-1}\left(\frac{\xi_{n-1}}{\lambda(t_{n-1})} + F(D_{n-1})\right),\ 1 \leq n \leq N.$$

We can write the cost function (7.13) as follows:

$$C^P(\boldsymbol{\xi},\boldsymbol{t}) = \sum_{n=0}^{N}\lambda(t_n)\int_{D_{t_n}}^{D_{t_n+}} xf(x)dx = \sum_{n=0}^{N}\lambda(t_n)\left[G\left(\frac{\lambda(t_n)F(D_n)+\xi_n}{\lambda(t_n)}\right) - G(F(D_n))\right]. \tag{7.51}$$

We begin with the following lemmas that we use to characterize the critical points of the optimization problem.

Lemma 7.5. *For* $i = 0,\ldots,N-1$, *we have the following equations:*

$$\frac{\partial C^P}{\partial \xi_i} = F^{-1}\left(\frac{\xi_i}{\lambda(t_i)} + F(D_i)\right) + \hat a_{i+1}\frac{f(D_{i+1})}{f\left(F^{-1}\left(\frac{\xi_i}{\lambda(t_i)}+F(D_i)\right)\right)}\left(\frac{\partial C^P}{\partial \xi_{i+1}} - D_{i+1}\right).$$

Proof. First, we have $\frac{\partial D_n}{\partial \xi_i} = 0$ for $i \geq n$, and the following recursive equations:

$$\frac{\partial D_n}{\partial \xi_{n-1}} = \frac{a_n}{\lambda(t_{n-1}) f\left(F^{-1}(\frac{\xi_{n-1}}{\lambda(t_{n-1})} + F(D_{n-1}))\right)}, \quad \frac{\partial D_n}{\partial \xi_i} = \frac{\hat{a}_{i+1} f(D_{i+1})}{f\left(F^{-1}(\frac{\xi_i}{\lambda(t_i)} + F(D_i))\right)} \frac{\partial D_n}{\partial \xi_{i+1}} \text{ for } 1 \leq i \leq n-2.$$

From (7.51), we get:

$$\begin{aligned}
\frac{\partial C^P}{\partial \xi_i} &= F^{-1}\left(\frac{\xi_i}{\lambda(t_i)} + F(D_i)\right) + \sum_{n=i+1}^{N} \left[F^{-1}\left(F(D_n) + \frac{\xi_n}{\lambda(t_n)}\right) - D_n\right] f(D_n) \frac{\partial D_n}{\partial \xi_i} \\
&= F^{-1}\left(\frac{\xi_i}{\lambda(t_i)} + F(D_i)\right) + \frac{\hat{a}_{i+1} f(D_{i+1})}{f\left(F^{-1}(\frac{\xi_i}{\lambda(t_i)} + F(D_i))\right)} \left[F^{-1}(F(D_{i+1}) + \frac{\xi_{i+1}}{\lambda(t_{i+1})}) - D_{i+1}\right] \\
&\quad + \frac{\hat{a}_{i+1} f(D_{i+1})}{f\left(F^{-1}(\frac{\xi_i}{\lambda(t_i)} + F(D_i))\right)} \left[\frac{\partial C^P}{\partial \xi_{i+1}} - F^{-1}\left(\frac{\xi_{i+1}}{\lambda(t_{i+1})} + F(D_{i+1})\right)\right],
\end{aligned}$$

which gives the result.

Lemma 7.6. *Under Assumption 7.2, we have that:*

1. *The function $x \mapsto x f(x)$ is increasing on $\mathbb{R}$ (or equivalently, $\tilde{F}$ is convex).*
2. *We have $f\left(\frac{x}{a_i}\right) - \hat{a}_i f(x) > 0, \quad i = 1, \ldots, N$.*
3. *The function*

$$x \in \mathbb{R}, \;\; h_{P,i}(x) = x \frac{\left[\frac{1}{a_i} f(\frac{x}{a_i}) - \hat{a}_i f(x)\right]}{f\left(\frac{x}{a_i}\right) - \hat{a}_i f(x)}$$

is well-defined, bijective increasing and satisfies $\operatorname{sgn}(x) h_{P,i}(x) \geq |x|$.

Proof. 1. We have $(x f(x))' > 0$ since $x f'(x) \geq 0$ by Assumption 7.2.
2. We have for $x \in \mathbb{R}$,

$$\lambda(t_{i-1}) f(\frac{x}{a_i}) - \lambda(t_i) a_i f(x) \geq \lambda(t_{i-1}) f(x)(1 - \hat{a}_i) > 0$$

because $f\left(\frac{x}{a_i}\right) \geq f(x)$ and $\hat{a}_i < 1$ by Assumption 7.2.
3. The function $h_{P,i}$ is well-defined thanks to the second point. We have $\operatorname{sgn}(x) h_{P,i}(x) \geq |x|$ since

$$h_{P,i}(x) = x \left[1 + \frac{a_i^{-1}}{1 - \hat{a}_i \frac{f(x)}{f(\frac{x}{a_i})}}\right],$$

and it is sufficient to check that $f(x)/f(x/a_i)$ is nondecreasing on $\mathbb{R}_+$ and nonincreasing on $\mathbb{R}_-$. We calculate

$$\left(\frac{f(x)}{f\left(\frac{x}{a_i}\right)}\right)' = \frac{f'(x) f\left(\frac{x}{a_i}\right) - \frac{1}{a_i} f(x) f'\left(\frac{x}{a_i}\right)}{f\left(\frac{x}{a_i}\right)^2}.$$

This is nonnegative on $\mathbb{R}_+$ and nonpositive on $\mathbb{R}_-$ if and only if $\frac{xf'(x)}{f(x)} \geq \frac{xf'(x/a_i)}{a_i f(x/a_i)}$ for $x \in \mathbb{R}$, which holds by Assumption 7.2 since $|x| \leq |x|/a_i$.

Proof of Theorem 7.5: We remark that the cost (7.51) can be written as follows:

$$C^P(\boldsymbol{\xi}, \boldsymbol{t}) = \lambda(t_N)\tilde{F}\left(F^{-1}\left(F(D_N) + \frac{\xi_N}{\lambda(t_N)}\right)\right) + \sum_{n=0}^{N-1} \lambda(t_n)\left[\tilde{F}\left(F^{-1}\left(F(D_n) + \frac{\xi_n}{\lambda(t_n)}\right)\right) - \frac{\lambda(t_{n+1})}{\lambda(t_n)}\tilde{F}\left(a_{n+1}F^{-1}\left(F(D_n) + \frac{\xi_n}{\lambda(t_n)}\right)\right)\right].$$

Since $\tilde{F}$ is convex by Lemma 7.6 and $\tilde{F}(0) = 0$, we have $\tilde{F}(a_{n+1}x) \leq a_{n+1}\tilde{F}(x)$, for $x \in \mathbb{R}$ and thus

$$C^P(\boldsymbol{\xi}, \boldsymbol{t}) \geq \lambda(t_N)\tilde{F}\left(F^{-1}\left(F(D_N) + \frac{\xi_N}{\lambda(t_N)}\right)\right) + \sum_{n=0}^{N-1} \lambda(t_n)\tilde{F}\left(F^{-1}\left(F(D_n) + \frac{\xi_n}{\lambda(t_n)}\right)\right)(1 - \hat{a}_{n+1}).$$

In particular $C^P(\boldsymbol{\xi}, \boldsymbol{t}) \geq 0$, since $\tilde{F} \geq 0$ and $\hat{a}_{n+1} < 1$ by Assumption (7.2). Besides, by setting $T(\boldsymbol{\xi}) = \left(\frac{\xi_0}{\lambda(t_0)}, D_1 + \frac{\xi_1}{\lambda(t_1)}, \ldots, D_N + \frac{\xi_N}{\lambda(t_N)}\right)$, we can easily check that $|T(\boldsymbol{\xi})| \underset{|\boldsymbol{\xi}| \to +\infty}{\to} +\infty$, which gives immediately that $C^P(\boldsymbol{\xi}, \boldsymbol{t}) \underset{|\boldsymbol{\xi}| \to +\infty}{\to} +\infty$ since $\tilde{F}(x) \underset{|x| \to +\infty}{\to} +\infty$.

Thus, there must be at least one minimizer of $C^P(\boldsymbol{\xi}, \boldsymbol{t})$ on $\{\boldsymbol{\xi} \in \mathbb{R}^{N+1}, \sum_{i=0}^N \xi_i = -\mathbb{x}\}$, and we denote by ν a Lagrange multiplier such that $\frac{\partial C^P}{\partial \xi_i} = \nu$. By Lemma 7.5 we obtain:

$$\nu = h_{P,i+1}(D_{i+1}), \quad i = 0, \ldots, N-1.$$

We also have $\frac{\partial C^P}{\partial \xi_N} = F^{-1}\left(F(D_N) + \frac{x_N}{\lambda(t_N)}\right) = \nu$, and we get $(i = 1, \ldots, N-1)$:

$$\xi_0^\star = \lambda(t_0)F\left(\frac{h_{P,1}^{-1}(\nu)}{a_1}\right), \; \xi_i^\star = \lambda(t_i)\left[F\left(\frac{h_{P,i+1}^{-1}(\nu)}{a_{i+1}}\right) - F\left(h_{P,i}^{-1}(\nu)\right)\right], \; \xi_N^\star = \lambda(t_N)\left[F(\nu) - F(h_{P,N}^{-1}(\nu))\right].$$

Besides, we have

$$\lambda(t_N)F(\nu) + \sum_{i=1}^N \lambda(t_{i-1})\left[F\left(\frac{h_{P,i}^{-1}(\nu)}{a_i}\right) - \frac{\lambda(t_i)}{\lambda(t_{i-1})}F(h_{P,i}^{-1}(\nu))\right] = -\mathbb{x}. \tag{7.52}$$

Since F is increasing bijective on $\mathbb{R}$ and the function $y \mapsto F(y) - \frac{\lambda(t_i)}{\lambda(t_{i-1})}F(a_i y)$ is increasing (its derivative is positive by Lemma 7.6), there is a unique solution to (7.52), and ν has the same sign as $-\mathbb{x}$. Thus $\boldsymbol{\xi}^\star$ is the unique optimal strategy. Moreover, the initial and the last trade have the same sign as $-\mathbb{x}$ since $\operatorname{sgn}(\nu)h_{P,N}(\nu) \geq |\nu|$. □

We now prepare the proof of Theorem 7.6. For sake of clearness, we will work under assumption (i) and assume that $\rho_t\left(1 + \frac{xf'(x)}{f(x)}\right) - \eta_t > 0$ for any $x \in \mathbb{R}$ and that $h_{P,t}$ is bijective and increasing.

However, a close look at the proof below is sufficient that the same arguments also work under assumption (ii).

Contrary to model V, it is more convenient to work with the process D rather than E (both are related by $D_t = F^{-1}(E_t/\lambda(t))$. We introduce for $0 \leq t \leq T$,

$$C^P(t,T,D_t,X_t) = \lambda(t)\left[G(\zeta_t) - \tilde{F}(D_t)\right] + \int_t^T \zeta_u \xi_u du + \lambda(T)[\tilde{F}(\nu) - G(\zeta_T)], \tag{7.53}$$

where

$$\nu \in \mathbb{R}, s.t. - E_t + \int_t^T \lambda(u)\left[\rho_u h_{P,u}^{-1}(\nu) f(h_{P,u}^{-1}(\nu)) - \eta_u F\left(h_{P,u}^{-1}(\nu)\right)\right] du + \lambda(T)F(\nu) = -X_t, \tag{7.54}$$

$$\zeta_u = h_{P,u}^{-1}(\nu),\ \xi_u = \lambda(u) f(\zeta_u)[\frac{d\zeta_u}{du} + \rho_u \zeta_u]. \tag{7.55}$$

Let us observe that $x \mapsto \rho_u x f(x) - \eta_u F(x)$ is increasing: its derivative is equal to $f(x)\left(\rho_u\left(1+\frac{xf'(x)}{f(x)}\right) - \eta_u\right)$ and is positive by assumption. Therefore, the left hand side of (7.54) is an increasing bijection on $\mathbb{R}$ and there is a unique solution ν to (7.54). The function $C^P(t,T,D_t,X_t)$ represents the minimal cost to liquidate X_t shares on $[t,T]$ given the current state D_t. We have in particular that $C^P(T,T,D_T,X_T) = \lambda(T)\left[G\left(\frac{E_T - X_T}{\lambda(T)}\right) - G\left(\frac{E_T}{\lambda(T)}\right)\right]$, which is the cost of selling X_T shares at time T. Besides, an integration by parts gives that

$$C^P(t,T,D_t,X_t) = -\lambda(t)\tilde{F}(D_t) + \int_t^T \lambda(u)\left[\rho_u f(\zeta_u)\zeta_u^2 - \eta_u \tilde{F}(\zeta_u)\right] du + \lambda(T)\tilde{F}(\nu). \tag{7.56}$$

The function $\zeta \mapsto \rho_u f(\zeta)\zeta^2 - \eta_u \tilde{F}(\zeta)$ is nonnegative: it vanishes for $\zeta = 0$ and its derivative is equal to $\zeta f(\zeta)\left(\rho_u\left(2+\frac{\zeta f'(\zeta)}{f(\zeta)}\right) - \eta_u\right)$ and has the same sign as ζ by assumption. Since $\tilde{F} \geq 0$, this gives

$$C^P(0,T,0,\mathbb{x}) \geq 0. \tag{7.57}$$

Proof of Theorem 7.6: Let $(X_t, 0 \leq t \leq T+)$ denote an admissible strategy that liquidates $\mathbb{x}$. We consider $(E_t, 0 \leq t \leq T+)$ the solution of $dE_t = dX_t + \eta_t E_t dt - \rho_t \lambda(t) f(F^{-1}(E_t/\lambda(t)))F^{-1}(E_t/\lambda(t))dt$, $D_t = F^{-1}(E_t/\lambda(t))$, ν_t the solution of (7.54) and $\zeta_t = h_{P,t}^{-1}(\nu_t)$. We set

$$C_t = \int_0^t D_s dX_s^c + \sum_{0 \leq s < t} \lambda(s)\left[G\left(\frac{E_s + \Delta X_s}{\lambda(s)}\right) - G\left(\frac{E_s}{\lambda(s)}\right)\right] + C^P(t,T,D_t,X_t).$$

Let us observe that $C_T = C^P(X)$ and $C_0 = C^P(0,T,0,\mathbb{x})$. We will show that $dC_t \geq 0$, and that $dC_t = 0$ holds only for $X^\star$. This will in particular prove that $C^P(X) \geq 0$ from (7.57).

Let us first consider the case of a jump $\Delta X_t > 0$. Then, we have

$$\Delta C_t = \lambda(t)\left[G\left(\frac{E_t + \Delta X_t}{\lambda(t)}\right) - G\left(\frac{E_t}{\lambda(t)}\right)\right] + C^P(t+,T,D_{t+},X_{t+}) - C^P(t,T,D_t,X_t).$$

Since $\Delta E_t = \Delta X_t$, we have $\nu_t = \nu_{t+}$ from (7.54) and then $\Delta C_t = 0$ since $\tilde{F}(D_t) = G(E_t/\lambda(t))$. Now, let us calculate dC_t. We set

$$\tilde{C}(t, T, D_t, X_t, v) = \lambda(T)\tilde{F}(v) - \lambda(t)\tilde{F}(D_t) + \int_t^T \lambda(u)\left[\rho_u f(h_{P,u}^{-1}(v))h_{P,u}^{-1}(v)^2 - \eta_u \tilde{F}(h_{P,u}^{-1}(v))\right] du.$$

Since $dD_t^c = -\rho_t D_t dt + \frac{dX_t^c}{\lambda(t)f(D_t)}$, we have from (7.56):

$$\begin{aligned} dC_t = {} & D_t dX_t^c - \lambda'(t)\tilde{F}(D_t)dt + \lambda(t)\rho_t f(D_t)D_t^2 dt - D_t dX_t^c - \lambda(t)[\rho_t f(\zeta_t)\zeta_t^2 - \eta_t \tilde{F}(\zeta_t)]dt \\ & + \frac{\partial \tilde{C}}{\partial v}(t, T, D_t, X_t, \nu_t)d\nu_t. \end{aligned}$$

Since $d(E_t - X_t) = \lambda(t)\left[\eta_t F(D_t) - \rho_t D_t f(D_t)\right] dt$, we get from (7.54)

$$\begin{aligned} & \left[\int_t^T \lambda(u)(h_{P,u}^{-1})'(\nu_t)\left[(\rho_u - \eta_u)f(h_{P,u}^{-1}(\nu_t)) + \rho_u h_{P,u}^{-1}(\nu_t)f'(h_{P,u}^{-1}(\nu_t))\right] du + \lambda(T)f(\nu_t)\right] d\nu_t \\ & \qquad - \lambda(t)\left[\rho_t h_{P,t}^{-1}(\nu_t)f(h_{P,t}^{-1}(\nu_t)) - \eta_t F\left(h_{P,t}^{-1}(\nu_t)\right)\right] dt = \lambda(t)\left[\eta_t F(D_t) - \rho_t D_t f(D_t)\right] dt. \end{aligned} \tag{7.58}$$

On the other hand, we have

$$\begin{aligned} \partial_v \tilde{C}(t, T, E_t, D_t, v) &= \lambda(T)vf(v) + \int_t^T \lambda(u)(h_{P,u}^{-1})'(v)h_{P,u}^{-1}(v)\left[(2\rho_u - \eta_u)f(h_{P,u}^{-1}(v)) + \rho_u h_{P,u}^{-1}(v)f'(h_{P,u}^{-1}(v))\right] du \\ &= v\left[\lambda(T)f(v) + \int_t^T \lambda(u)(h_{P,u}^{-1})'(v)\left((\rho_u - \eta_u)f(h_{P,u}^{-1}(v)) + \rho_u h_{P,u}^{-1}(v)f'(h_{P,u}^{-1}(v))\right) du\right], \end{aligned}$$

and we get $\frac{\partial \tilde{C}}{\partial v}(t, T, D_t, X_t, \nu_t)d\nu_t = \lambda(t)\nu_t[\eta_t(F(D_t) - F(\zeta_t)) + \rho_t(\zeta_t f(\zeta_t) - D_t f(D_t))]$. We finally obtain:

$$\begin{aligned} & dC_t = \lambda(t)\psi_t(\zeta_t)dt, \text{ with} \\ & \psi_t(\zeta) = \eta_t(\tilde{F}(\zeta) - \tilde{F}(D_t)) + \rho_t(D_t^2 f(D_t) - \zeta^2 f(\zeta)) + h_{P,t}(\zeta)\left(\eta_t(F(D_t) - F(\zeta)) + \rho_t(\zeta f(\zeta) - D_t f(D_t))\right). \end{aligned} \tag{7.59}$$

We have $\psi_t(D_t) = 0$ and get that $\psi_t'(\zeta) = h_{P,t}'(\zeta)\left[\eta_t(F(D_t) - F(\zeta)) + \rho_t(\zeta f(\zeta) - D_t f(D_t))\right]$ by simple calculations. On the one hand, we have $h_{P,t}'(\zeta) > 0$. On the other hand, the bracket is positive on $\zeta > D_t$ and negative on $\zeta < D_t$ since its derivative is equal to $(\rho_t - \eta_t)f(\zeta) + \rho_t \zeta f(\zeta)$, which is positive by assumption. Thus, D_t is the unique minimum of ψ_t: $\psi_t(D_t) = 0$ and $\psi_t(\zeta) > 0$ for $\zeta \neq D_t$.

Thus, if X is an optimal strategy, we necessarily have $\zeta_t = D_t$, dt-a.e. From (7.58), we get

$$\left[\int_t^T \lambda(u)(h_{P,u}^{-1})'(\nu_t)\left[(\rho_u - \eta_u)f(h_{P,u}^{-1}(\nu_t)) + \rho_u h_{P,u}^{-1}(\nu_t)f'(h_{P,u}^{-1}(\nu_t))\right] du + \lambda(T)f(\nu_t)\right] d\nu_t = 0,$$

and thus $d\nu_t = 0$ since $(h_{P,u}^{-1})'$ and $x \mapsto (\rho_u - \eta_u)f(x) + \rho_u x f'(x)$ are positive functions by assumption. We get that $\nu_t = \nu$, where ν is the solution of (7.38). In particular, we have $\Delta X_0 = \lambda(0)F(D_{0+}) = \lambda(0)F(h_{P,0}^{-1}(\nu)) = \Delta X_0^\star$ and then $X = X^\star$. This gives the uniqueness of the optimal strategy. Last, $\xi_0^\star$ has the same sign as $-\mathbb{x}$ since ν and $-\mathbb{x}$ have the same sign. $\square$

Proof of Corollary 7.6: By Assumption 7.2 we have $\rho_t - \eta_t > 0$, $xf'(x) \geq 0$ and $x\partial_x(\frac{xf'(x)}{f(x)}) \leq 0$, which gives:

$$h'_{P,t}(x) = \frac{\left(\rho_t\left(2+\frac{xf'(x)}{f(x)}\right)-\eta_t\right)\left(\rho_t\left(1+\frac{xf'(x)}{f(x)}\right)-\eta_t\right)-\rho_t^2 x\partial_x(\frac{xf'(x)}{f(x)})}{\left(\rho_t\left(1+\frac{xf'(x)}{f(x)}\right)-\eta_t\right)^2} > 0.$$

Also, we have $\mathrm{sgn}(x)h_{P,t}(x) \geq |x|$, and $h_{P,t}$ is thus bijective on $\mathbb{R}$. We deduce that $\mathrm{sgn}(x)h^{-1}_{P,t}(x) \leq |x|$, which gives that the last trade $\xi^\star_T$ has the same sign as $-\mathbb{x}$.

Let us assume moreover that (7.30) holds. Let $\gamma_t = \frac{\lambda(t)f(\zeta_t)\zeta_t}{h'_{P,t}(\zeta_t)\left(1+\frac{\zeta_t f'(\zeta_t)}{f(\zeta_t)}-\frac{\eta_t}{\rho_t}\right)^2} > 0$. Then,

$$\begin{aligned}
\xi_t &= \gamma_t\left[\frac{\rho'_t\eta_t-\rho_t\eta'_t}{\rho_t^2}+\rho_t\left(1+\frac{\zeta_t f'(\zeta_t)}{f(\zeta_t)}-\frac{\eta_t}{\rho_t}\right)\left(2+\frac{\zeta_t f'(\zeta_t)}{f(\zeta_t)}-\frac{\eta_t}{\rho_t}\right)-\zeta_t\partial_x(\frac{xf'(x)}{f(x)})|_{x=\zeta_t}\right]\\
&\geq \gamma_t\left[\frac{\rho'_t\eta_t-\rho_t\eta'_t}{\rho_t^2}+\rho_t\left(1-\frac{\eta_t}{\rho_t}\right)\left(2-\frac{\eta_t}{\rho_t}\right)\right] \text{ by Assumption 7.2.}\\
&= \gamma_t\left(\frac{2\rho_t-\eta_t}{\rho_t}\right)^2\left[\left(\frac{\rho_t-\eta_t}{2\rho_t-\eta_t}\right)'+\rho_t\left(\frac{\rho_t-\eta_t}{2\rho_t-\eta_t}\right)\right] \geq 0 \text{ by (7.30).}
\end{aligned}$$

□

References

[AP05] Y. Achdou and O. Pironneau. *Computational methods for option pricing*, volume 30 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[AFS08] A. Alfonsi, A. Fruth, and A. Schied. Constrained portfolio liquidation in a limit order book model. In *Advances in mathematics of finance*, volume 83 of *Banach Center Publ.*, pages 9–25. Polish Acad. Sci. Inst. Math., Warsaw, 2008.

[AFS10] A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2):143–157, 2010.

[AS10] A. Alfonsi and A. Schied. Optimal trade execution and absence of price manipulations in limit order book models. *SIAM Journal on Financial Mathematics*, 1(1):490–522, 2010.

[ASS11] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation and the positive portfolio problem. *SSRN eLibrary*, 2011.

[Alm03] R. F. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10(1):1–18, 2003.

[AC00] R. F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.

[AHLT05] R. F. Almgren, E. Hauptmann, H. Li, and C. Thum. Direct estimation of equity market impact. *Risk*, 18:57–62, jul 2005.

[AL07] R. F. Almgren and J. Lorenz. Adaptive arrival price. *Algorithmic Trading III: Precision, Control, Execution*, 2007.

[ACKM06] A. Ammar, F. Chinesta, R. Keunings, and B. Mokdad. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *J. Non-Newtonian Fluid Mech.*, 139:153–176, 2006.

[ADM97] M. Avellaneda, G. Davis, and S. Mallat. Adaptive greedy approximations. *Constr. Approx.*, 13(1):57–98, 1997.

[AS08] M. Avellaneda and S. Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.

[Bacal] L. Bachelier. *Théorie de la spéculation*. Les grandes classiques Gauthiers-Villars. Éditions Jacques Gabay, Sceaux, 1995, Reprint of the 1900 original.

[BK09] B. W. Bader and T. G. Kolda. Tensor decompositions and applications. *SIAM Rev.*, 51(3):455–500, 2009.

[BF12] P. Bank and A. Fruth. Optimal order scheduling for deterministic liquidity patterns. *Preprint*, 2012.

[BCDD08] A. Barron, A. Cohen, W. Dahmen, and R. DeVore. Approximation and learning by greedy algorithms. *Annals of Statistics*, 36(1):64–94, 2008.

[BL98] D. Bertsimas and A. W. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1 – 50, 1998.

[BM02] G. Beylkin and M. J. Mohlenkamp. Numerical operator calculus in higher dimensions. *Proc. Natl. Acad. Sci. USA*, 99(16):10246–10251 (electronic), 2002.

[BHS95] B. Biais, P. Hillion, and C. Spatt. An empirical analysis of the limit order book and the order flow in the Paris Bourse. *The Journal of Finance*, 50(5):1655–1689, 1995.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):pp. 637–654, 1973.

[BDdL+11] J.-P. Bouchaud, C. Deremble, J. de Lataillade, J. Kockelkoren, Y. Lempérière, and B. Tóth. Anomalous price impact and the critical nature of liquidity in financial markets. *Phys. Rev. X*, 1:021006, Oct 2011.

[BGPW04] J.-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of «random»price changes. *Quantitative Finance*, 4(2):176–190, 2004.

[BP03] J.-P. Bouchaud and M. Potters. More statistical properties of order books and price impact. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):133–140, 2003. Proceedings of the International Econophysics Conference.

[BLM09] C. Le Bris, T. Lelièvre, and Y. Maday. Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations. *Constructive Approximation*, 30(3):621–651, 2009.

[BP05] M. K. Brunnermeier and L. H. Pedersen. Predatory trading. *The Journal of Finance*, 60(4):1825–1863, 2005.

[BG04] H. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.

[CEL12] E. Cancès, V. Ehrlacher, and T. Lelièvre. Convergence of a greedy algorithm for high-dimensional convex nonlinear problems. 2012. Mathematical Models and Methods in Applied Sciences.

[CLV07] B. I. Carlin, M. S. Lobo, and S. Viswanathan. Episodic liquidity crises: Cooperative and predatory trading. *The Journal of Finance*, 62(5):2235–2274, 2007.

[CMS99] P. Carr, D. B. Madan, and R. H. Smith. Option valuation using the Fast Fourier Transform. *Journal of Computational Finance*, 2:61–73, 1999.

[ÇJP04] U. Çetin, R. A. Jarrow, and P. Protter. Liquidity risk and arbitrage pricing theory. *Finance Stoch.*, 8(3):311–341, 2004.

[CRS01] T. Chordia, R. Roll, and A. Subrahmanyam. Market liquidity and trading activity. *The Journal of Finance*, 56(2):501–530, 2001.

[CD11] R. Cont and A. De Larrard. Price dynamics in a Markovian limit order market. *Social Science Research Network Working Paper Series*, January 2011.

[dSL08] V. de Silva and L. Lim. Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1084–1127, 2008.

[DHM89] R. DeVore, R. Howard, and C. Micchelli. Optimal nonlinear approximation. *Manuscripta mathematica*, 63(4):469–478, 1989.

[DT96] R. A. DeVore and V. N. Temlyakov. Some remarks on greedy algorithms. *Advances in Computational Mathematics*, 5:173–187, 1996.

[DFS03] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.

[DPS00] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376, 2000.

[ES09] L. Eldén and B. Savas. A Newton-Grassmann method for computing the best multilinear rank-(r_1, r_2, r_3). approximation of a tensor. *SIAM Journal on Matrix Analysis and Applications*, 31:248–271, 2009.

[FGL12] J. Fernandez, O. Guéant, and C. Lehalle. Optimal portfolio liquidation with limit orders. *SIAM Journal on Financial Mathematics*, 3(1):740–764, 2012.

[FS12] L. Figueroa and E. Suli. Greedy approximation of high-dimensional Ornstein-Uhlenbeck operators with unbounded drift. *Foundations of Computational Mathematics*, 12(5):573–623, 2012.

[FSU13] A. Fruth, T. Schöneborn, and M. Urusov. Optimal trade execution and price manipulation in order books with time-varying liquidity. *Mathematical Finance*, 2013.

[Gat10] J. Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759, 2010.

[GSS12] J. Gatheral, A. Schied, and A. Slynko. Transient linear price impact and Fredholm integral equations. *Math. Finance*, 22(3):445–474, 2012.

[GM92] M.S. Gerety and J. H. Mulherin. Trading halts and market activity: An analysis of volume at the open and the close. *The Journal of Finance*, 47(5):1765–1784, 1992.

[HH95] Y. Hamao and J. Hasbrouck. Securities trading in the absence of dealers: trades and quotes on the Tokyo Stock Exchange. *Review of Financial Studies*, 8(3):849–878, 1995.

[HS04] G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275, 2004.

[JJ88] C. P. Jain and G.-H. Joh. The dependence between hourly prices and trading volume. *The Journal of Financial and Quantitative Analysis*, 23(3):pp. 269–283, 1988.

[KM08] A. Kempf and D. Mayston. Liquidity commonality beyond best prices. *Journal of Financial Research*, 31(1):25–40, 2008.

[KW96] A. W. Kleidon and I. M. Werner. U.K. and U.S. trading of British cross-listed stocks: an intraday analysis of market integration. *Review of Financial Studies*, 9(2):619–664, 1996.

[KSS11] F. Klöck, A. Schied, and Y. Sun. Price manipulation in a market impact model with dark pool. *Social Science Research Network Working Paper Series*, August 2011.

[KT99] S. V. Konyagin and V. N. Temlyakov. Rate of convergence of pure greedy algorithm. *East Journal on Approximations*, 5:493–499, 1999.

[KL09] O. Kudryavtsev and S. Levendorskii. Fast and accurate pricing of barrier options under lévy processes. *Finance and Stochastics*, 13(4):531–562, 2009.

[Lad99] P. Ladevèze. Nonlinear computational structural mechanics: new approaches and non-incremental methods of calculations. 1999.

[LMV00a] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21:1253–1278, 2000.

[LMV00b] L. De Lathauwer, B. De Moor, and J. Vandewalle. On the best rank-1 and rank-$(r_1, r_2, \ldots, r_n)$ approximation and applications of higher-order tensors. *SIAM Journal on Matrix Analysis and Applications*, 21:1324–1342, 2000.

[LO08] C. C. W. Leentvaar and C.W. Oosterlee. Multi-Asset Option Pricing Using A Parallel Fourier-Based Technique. *Journal of Computational Finance*, 12(1):1 – 26, October 2008.

[LO09] J. Lorenz and J. Osterrieder. Simulation of a limit order driven market. *The Journal Of Trading*, 4:23–30, 2009.

[Mer76] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125 – 144, 1976.

[Nou10] A. Nouy. Proper generalized decompositions for a priori model reduction of problems formulated in tensor product spaces: Alternative definitions and algorithms. *Proceedings of the Seventh International Conference on Engineering Computational Technology, Civil-Comp Press*, 2010.

[OW05] A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Preprint*, 2005.

[PS04] T. Von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *M2AN Math. Model. Numer. Anal.*, 38(1):93–127, 2004.

[Pom08] D. Pommier. *Méthodes numériques sur des grilles sparse appliquées à l'évaluation d'options en finance*. PhD thesis, Université Pierre et Marie Curie, 2008.

[PSS11] S. Predoiu, G. Shaikhet, and S. Shreve. Optimal execution in a general one-sided limit-order book. *SIAM J. Financial Math.*, 2:183–212, 2011.

[RW05] B. Rosenow and P. Weber. Order book approach to price impact. *Quantitative Finance*, 5(4):357–364, 2005.

[Sch13] A. Schied. Robust strategies for optimal order execution in the Almgren-Chriss framework. *Applied Mathematical Finance*, 20(3):264–286, 2013.

[SS09] A. Schied and T. Schöneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastics*, 13:181–204, 2009.

[Smo63] S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Mathematics Doklady*, 3:240–243, 1963.

[Tem08] V. N. Temlyakov. Greedy approximation. *Acta Numerica*, 17:235–409, 2008.

[Zen91] C. Zenger. Sparse grids. In *Parallel algorithms for partial differential equations (Kiel, 1990)*, volume 31 of *Notes Numer. Fluid Mech.*, pages 241–251. Vieweg, Braunschweig, 1991.